Midterm 2 Review

Natalie Parde, Ph.D.

Department of Computer Science

University of Illinois at Chicago

CS 421: Natural Language Processing

Fall 2019

Exam Format

True/False (30%)

- 15 questions, 2 points each
- No partial credit (autograded)

Multiple Choice (40%)

- 8 questions, 5 points each
- No partial credit (autograded)

Problem Solving (30%)

- 3 questions, 10 points each
- Show work for partial credit (graded manually)

Bonus Question (10%)

- Problem-solving question
- Points added to exam score (max exam score = 100)

Sample Midterm

- Currently available on Piazza
- Solution will be posted after class

What should I study?

- Questions designed based on **slides** and **assignments**
- Problem solving questions similar but not identical to **problem solving examples** (e.g., stepping through a neural network) in slides

What content will the exam cover?

- N-Grams
- Language Modeling
- Word Embeddings
- Cosine Similarity
- Naive Bayes
- Text Classification
- Evaluation Metrics
- Feedforward Neural Networks
- Recurrent Neural Networks

What will I for sure *not* need to memorize?

- Derivatives
- Log values
- Language model interpolation
- Word embedding visualization techniques
- Implementation details for GloVe, ELMo, or BERT
- Backpropagation equations
    - Do have a general sense of how backpropagation works
- Implementation details for LSTM and GRU gates

What should I bring to the exam?

- Pen or pencil
- UIN (you'll need to write it on the first page of the exam)

- This exam will be:
  - Closed note
  - Closed book
  - Closed device
- You will not need a calculator

How long will the exam last?

Full class period, if needed (75 minutes)

True/False

Multiple Choice

Solution Time!

	W1	W2	W3
pumpkin	2	1	2
candy	2	2	1
halloween	1	2	1

Problem-Solving Questions

Consider the following word embeddings for *pumpkin*, *candy*, and *halloween*.

Compute the cosine similarities between (a) *pumpkin* and *halloween*, and (b) *candy* and *halloween*, to determine whether *pumpkin* or *candy* is closer to *halloween*. If you want, you can leave the final values as fractions rather than converting them to decimals.

	W1	W2	W3
pumpkin	2	1	2
candy	2	2	1
halloween	1	2	1

$$\cos(x,y) = \frac{x \cdot y}{|x||y|} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}}$$

Problem-Solving Questions

Consider the following word embeddings for *pumpkin*, *candy*, and *halloween*.

Compute the cosine similarities between (a) *pumpkin* and *halloween*, and (b) *candy* and *halloween*, to determine whether *pumpkin* or *candy* is closer to *halloween*. If you want, you can leave the final values as fractions rather than converting them to decimals.

	W1	W2	W3
pumpkin	2	1	2
candy	2	2	1
halloween	1	2	1

$$\cos(x,y) = \frac{x \cdot y}{|x||y|} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}}$$

$$\cos(\text{pumpkin, halloween}) = \frac{2*1+1*2+2*1}{\sqrt{2^2+1^2+2^2}\sqrt{1^2+2^2+1^2}}$$

Problem-Solving Questions

Consider the following word embeddings for *pumpkin*, *candy*, and *halloween*.

Compute the cosine similarities between (a) *pumpkin* and *halloween*, and (b) *candy* and *halloween*, to determine whether *pumpkin* or *candy* is closer to *halloween*. If you want, you can leave the final values as fractions rather than converting them to decimals.

	W1	W2	W3
pumpkin	2	1	2
candy	2	2	1
halloween	1	2	1

$$\cos(x,y) = \frac{x \cdot y}{|x||y|} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}}$$

$$\cos(\text{pumpkin, halloween}) = \frac{2*1+1*2+2*1}{\sqrt{2^2+1^2+2^2}\sqrt{1^2+2^2+1^2}} = \frac{6}{\sqrt{9}*\sqrt{6}}$$

Problem-Solving Questions

Consider the following word embeddings for *pumpkin*, *candy*, and *halloween*.

Compute the cosine similarities between (a) *pumpkin* and *halloween*, and (b) *candy* and *halloween*, to determine whether *pumpkin* or *candy* is closer to *halloween*. If you want, you can leave the final values as fractions rather than converting them to decimals.

	W1	W2	W3
pumpkin	2	1	2
candy	2	2	1
halloween	1	2	1

$$\cos(x,y) = \frac{x \cdot y}{|x||y|} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}}$$

$$\cos(\text{pumpkin, halloween}) = \frac{2*1+1*2+2*1}{\sqrt{2^2+1^2+2^2}\sqrt{1^2+2^2+1^2}} = \frac{6}{\sqrt{9}*\sqrt{6}}$$

$$\cos(\text{candy, halloween}) = \frac{2*1+2*2+1*1}{\sqrt{2^2+2^2+1^2}\sqrt{1^2+2^2+1^2}} = \frac{7}{\sqrt{9}*\sqrt{6}}$$

Problem-Solving Questions

Consider the following word embeddings for *pumpkin*, *candy*, and *halloween*.

Compute the cosine similarities between (a) *pumpkin* and *halloween*, and (b) *candy* and *halloween*, to determine whether *pumpkin* or *candy* is closer to *halloween*. If you want, you can leave the final values as fractions rather than converting them to decimals.

	W1	W2	W3
pumpkin	2	1	2
candy	2	2	1
halloween	1	2	1

$$\cos(x,y) = \frac{x \cdot y}{|x||y|} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2}\sqrt{\sum_{i=1}^{N} y_i^2}}$$

$$\cos(\text{pumpkin, halloween}) = \frac{2*1+1*2+2*1}{\sqrt{2^2+1^2+2^2}\sqrt{1^2+2^2+1^2}} = \frac{6}{\sqrt{9}*\sqrt{6}}$$

$$\cos(\text{candy, halloween}) = \frac{2*1+2*2+1*1}{\sqrt{2^2+2^2+1^2}\sqrt{1^2+2^2+1^2}} = \frac{7}{\sqrt{9}*\sqrt{6}}$$

Problem-Solving Questions

Consider the following word embeddings for *pumpkin*, *candy*, and *halloween*.

Compute the cosine similarities between (a) *pumpkin* and *halloween*, and (b) *candy* and *halloween*, to determine whether *pumpkin* or *candy* is closer to *halloween*. If you want, you can leave the final values as fractions rather than converting them to decimals.

document
Happy Halloween!
Tomorrow, Halloween candy will be on sale in the candy aisle.
Do you prefer Thanksgiving?

Problem Solving Questions

- Given the following set of documents, compute the TF*IDF for:
  1. Halloween, for the first document
  2. candy, for the second document
- Compute *tf* and *idf* without applying any smoothing operations or log functions. If you want, you can leave the final values as fractions rather than converting them to decimals.

Problem Solving Questions

document
Happy Halloween!
Tomorrow, Halloween candy will be on sale in the candy aisle.
Do you prefer Thanksgiving?

TF*IDF = count(term, doc) * (# docs) / (# docs w/term)

- Given the following set of documents, compute the TF*IDF for:
  1. Halloween, for the first document
  2. candy, for the second document
- Compute *tf* and *idf* without applying any smoothing operations or log functions. If you want, you can leave the final values as fractions rather than converting them to decimals.

document
Happy Halloween!
Tomorrow, Halloween candy will be on sale in the candy aisle.
Do you prefer Thanksgiving?

TF*IDF = count(term, doc) * (# docs) / (# docs w/term)

TF(Halloween, doc1) = 1

Problem Solving Questions

- Given the following set of documents, compute the TF*IDF for:
  1. Halloween, for the first document
  2. candy, for the second document
- Compute *tf* and *idf* without applying any smoothing operations or log functions. If you want, you can leave the final values as fractions rather than converting them to decimals.

document
Happy Halloween!
Tomorrow, Halloween candy will be on sale in the candy aisle.
Do you prefer Thanksgiving?

TF*IDF = count(term, doc) * (# docs) / (# docs w/term)

TF(Halloween, doc1) = 1
IDF(Halloween) = 3/2

Problem Solving Questions

- Given the following set of documents, compute the TF*IDF for:
  1. Halloween, for the first document
  2. candy, for the second document
- Compute *tf* and *idf* without applying any smoothing operations or log functions. If you want, you can leave the final values as fractions rather than converting them to decimals.

document
Happy Halloween!
Tomorrow, Halloween candy will be on sale in the candy aisle.
Do you prefer Thanksgiving?

TF*IDF = count(term, doc) * (# docs) / (# docs w/term)

TF(Halloween, doc1) = 1
IDF(Halloween) = 3/2
TF*IDF(Halloween, doc1) = 1 * 3/2 = 3/2

Problem Solving Questions

- Given the following set of documents, compute the TF*IDF for:
    1. Halloween, for the first document
    2. candy, for the second document
- Compute *tf* and *idf* without applying any smoothing operations or log functions. If you want, you can leave the final values as fractions rather than converting them to decimals.

Problem Solving Questions

- Given the following set of documents, compute the TF*IDF for:
  1. Halloween, for the first document
  2. candy, for the second document
- Compute *tf* and *idf* without applying any smoothing operations or log functions. If you want, you can leave the final values as fractions rather than converting them to decimals.

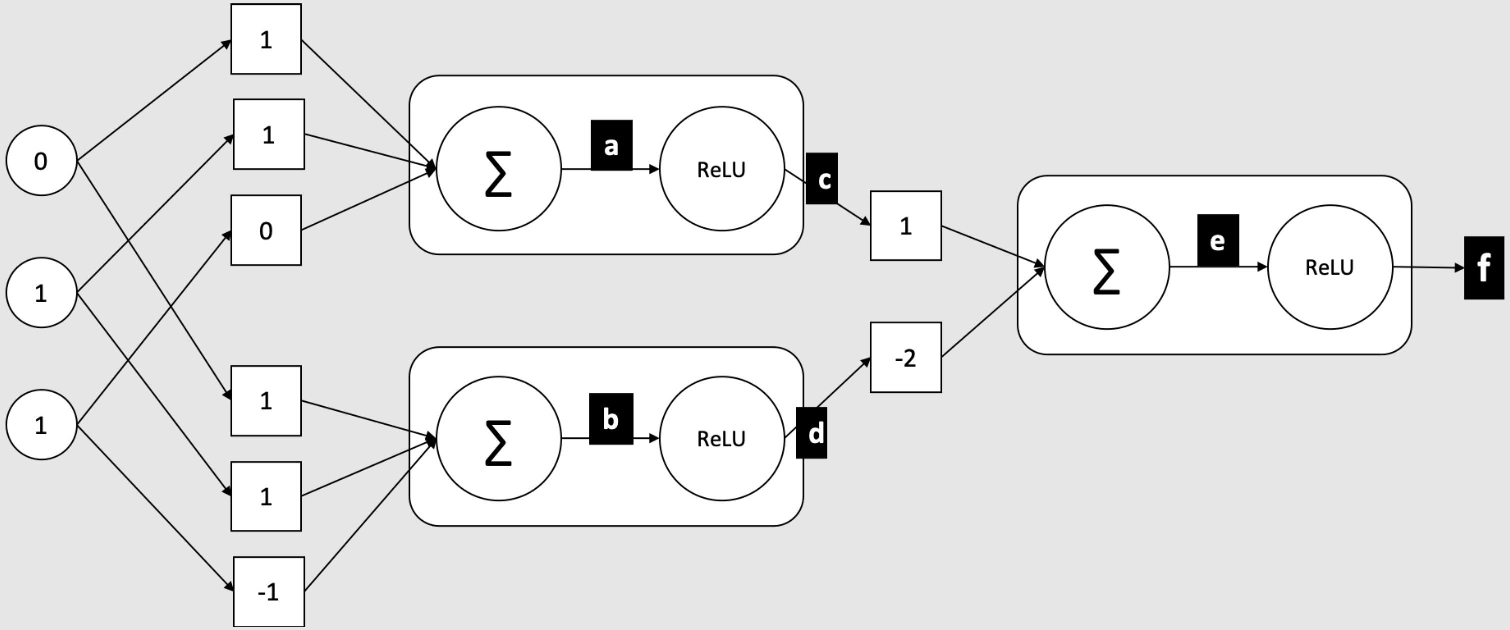
document
Happy Halloween!
Tomorrow, Halloween candy will be on sale in the candy aisle.
Do you prefer Thanksgiving?

TF*IDF = count(term, doc) * (# docs) / (# docs w/term)

TF(Halloween, doc1) = 1
IDF(Halloween) = 3/2
TF*IDF(Halloween, doc1) = 1 * 3/2 = 3/2

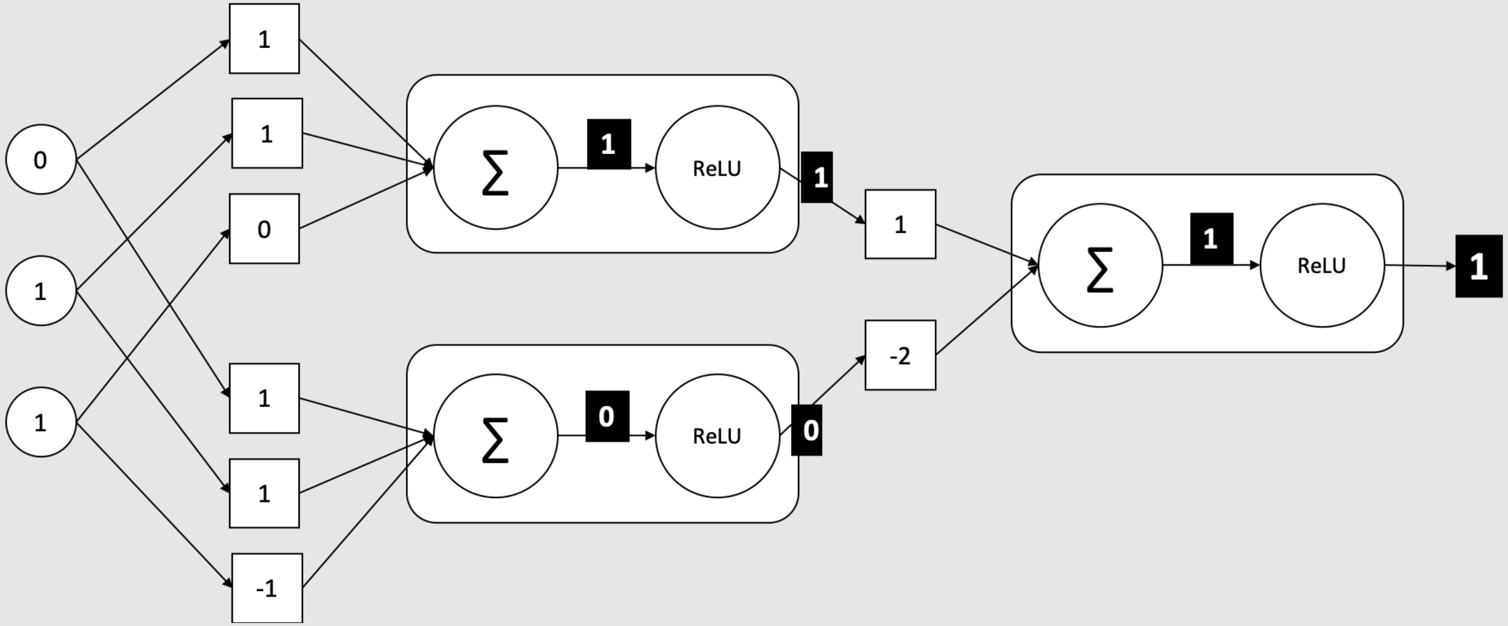
TF(candy, doc2) = 2
IDF(candy) = 3/1
TF*IDF(candy, doc2) = 2 * 3/1 = 6/1

Problem Solving Questions



- Given the feedforward neural network with the specified inputs, weights, and activation functions below, compute the network's intermediate (or final) values at the positions *a*, *b*, *c*, *d*, *e*, and *f* of the network that are specified.

Problem Solving Questions



- Given the feedforward neural network with the specified inputs, weights, and activation functions below, compute the network's intermediate (or final) values at the positions *a*, *b*, *c*, *d*, *e*, and *f* of the network that are specified.

Bonus Question

- Given the following contingency table, compute the microaveraged precision across all classes. If you want, you can leave the final values as fractions rather than converting them to decimals.

	Actual A	Actual B	Actual C
Predicted A	1	2	1
Predicted B	2	2	1
Predicted C	2	1	2

Bonus Question

- Given the following contingency table, compute the microaveraged precision across all classes. If you want, you can leave the final values as fractions rather than converting them to decimals.

	Actual A	Actual B	Actual C
Predicted A	1	2	1
Predicted B	2	2	1
Predicted C	2	1	2

	Actual A	Actual Not A
Predicted A	1	3
Predicted Not A	4	6

Bonus Question

- Given the following contingency table, compute the microaveraged precision across all classes. If you want, you can leave the final values as fractions rather than converting them to decimals.

	Actual A	Actual B	Actual C
Predicted A	1	2	1
Predicted B	2	2	1
Predicted C	2	1	2

	Actual A	Actual Not A
Predicted A	1	3
Predicted Not A	4	6

	Actual B	Actual Not B
Predicted B	2	3
Predicted Not B	3	6

Bonus Question

- Given the following contingency table, compute the microaveraged precision across all classes. If you want, you can leave the final values as fractions rather than converting them to decimals.

	Actual A	Actual B	Actual C
Predicted A	1	2	1
Predicted B	2	2	1
Predicted C	2	1	2

	Actual A	Actual Not A
Predicted A	1	3
Predicted Not A	4	6

	Actual B	Actual Not B
Predicted B	2	3
Predicted Not B	3	6

	Actual C	Actual Not C
Predicted C	2	3
Predicted Not C	2	7

Bonus Question

- Given the following contingency table, compute the microaveraged precision across all classes. If you want, you can leave the final values as fractions rather than converting them to decimals.

	Actual A	Actual B	Actual C
Predicted A	1	2	1
Predicted B	2	2	1
Predicted C	2	1	2

	Actual A	Actual Not A
Predicted A	1	3
Predicted Not A	4	6

	Actual B	Actual Not B
Predicted B	2	3
Predicted Not B	3	6

	Actual C	Actual Not C
Predicted C	2	3
Predicted Not C	2	7

	Actual x	Actual Not x
Predicted x	5	9
Predicted Not x	9	19

Bonus Question

- Given the following contingency table, compute the microaveraged precision across all classes. If you want, you can leave the final values as fractions rather than converting them to decimals.

	Actual A	Actual B	Actual C
Predicted A	1	2	1
Predicted B	2	2	1
Predicted C	2	1	2

	Actual A	Actual Not A
Predicted A	1	3
Predicted Not A	4	6

	Actual B	Actual Not B
Predicted B	2	3
Predicted Not B	3	6

	Actual C	Actual Not C
Predicted C	2	3
Predicted Not C	2	7

	Actual x	Actual Not x
Predicted x	5	9
Predicted Not x	9	19

Precision = TP / (TP+FP)
= 5 / (5+9) = 5/14

A few remaining details….

Exam Location: Same classroom as always (TBH 180G)

Exam Time: Same time as class (Tuesday from 9:30-10:45 a.m.)

Good luck!