Improving Machine Recognition of Collaborative Dialog Acts via Sentence Embeddings Students: Justice Parham, Vondarrian Patrick Advisor: Dr. Jung Hee Kim Department of Computer Science North Carolina Agricultural & Technical State University

Abstract

Introduction

This project attempts to train a computer to recognize linguistic dialogue acts within transcripts of students working together. In COMPS (Computer-Mediated **Problems Solving) exercises students work together** via typed-chat, solving problems in small groups in a computer science class.

| use refresh or back to change pages while o | n this site |
|---|--|
| | Group 11 |
| Hover to Show Previous Answ 1: 12 | wer Student A: The Typed of the exception does not really matter in this problem because the catch block covers all types of exception 14:31:50.308 |
| 2: 56 | Student B: so it would be 4 14:32:28.221 |
| 4: 56 5: 100 | Student C: don't we need to go through the finally block? 14:33:08.647 |
| | Student A: I think the program outputs the first couple of lines before the code crash before go through the catch and finally blocks 14:34:38.726 |
| | Student B: ok so 1: 12 14:35:32.376 |
| | Student B: 2: 56, 4: , 5:100? 14:35:49.452 |
| | Student C: I agree with you, but i think 56 should be in front of 4 14:37:02.953 |
| | Student A: I think it also print 4: 56 14:37:28.485 |
| | Student B: Why wouldn't 56 go with 2 14:38:08.470 |
| Answer is unlocked | Student A: It goes with both 14:39:05.829 |
| Student B is sequesting submi | Student C: I agree with A 14:39:15.314 |
| Student B is requesting submi | |
| Student C agrees | Student A: Yes 14:40:04.001 |
| | Student C: I believe so 14:40:31.724 |
| | Student B: Let lock in the answer and wait for TA to check 14:40:57.228 |
| Student A agrees | |

Figure 1: Chat Engineer Interface

Student dialogue turns are classified according to four categories of collaborative utterance: sharing ideas, negotiating ideas, regulating problem-solving, and maintaining communication.

Methods

- Pre-process dialogue text into set of numbers for each dialogue turn, then train a linear classifier.
- Atempt to improve classifier accuracy by changing the way sentences are converted into numerical features
- Use Doc2Vec, which is word order sensitive, in place of Topic modeling
- Utilize dialogue context by presenting preceding dialogue turns to the classifier

Goal

This research advances toward promoting better student collaborative problem-solving exercises, fully using student group cognition and more collaboration skills, and developing computer monitoring of the student conversation groups.

Annotation

| Figure 2: Student Dialogue Act Categories | | | | | | Line |
|--|----------------|--|---|---|------------------|-------------------|
| Category | Category Label | Category Description | Categorization Student groups worked together via typed-chat The dialogue turns were then categorized according to the 4 dialogue act categories. 715 hand-categorized dialogue turns were used in this experiment | | | |
| Sharing Ideas | A | Student shares task relevant idea that contributes information to the process. | | | | 0.6 - 0.4 - |
| Negotiating Ideas | B | The student listens and responds to a previously stated idea. They will either agree or disagree. | | | | 0.2 - |
| Regulation of C Problem | | The student attempts to direct or regulate | Figure 3: Transcript Categorization | | | Fig |
| Solving | | workflow. | Participa | nt Text | Dialogue Acts | Sei |
| Maintaining Communication | D | The participant contributes something that is not task relevant to | St1 | <pre>public String toStrong(){ String result = null; result = lendingInstitution +' '+ PAmount +' '+ iRate +' '+ etc.</pre> | A | Two wo ma |
| | | the group. | St2 | lol yall going in i think thats right tho | D, B | abl |
| Figure 3 demonstrates how student dialogues are analyzed and categorized into one of the four categories. Some turns show attributes of more than one category and are labeled accordingly. | | St1 | we just have to explain the getters and setters now | С | inte Res | |
| | | St3 | Student 1 can u explain them | С | • It a | |
| | | St1 | besides excapsulation, accessors make it easier to change future things mybad on the spelling | A, D | cat 0.3 | |
| | | St4 | So everything except the setters and getters are explained right? | С | | |

Topic Modeling Result

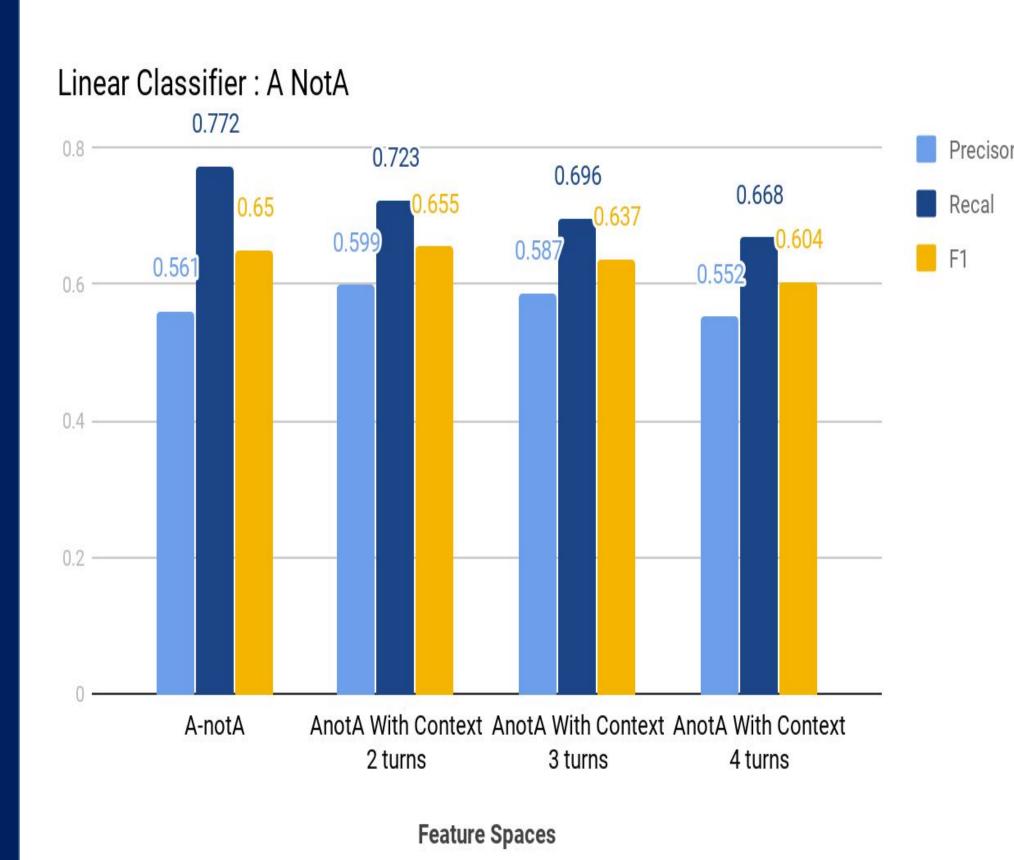


Figure 4: Topic Model results

OMPS

several

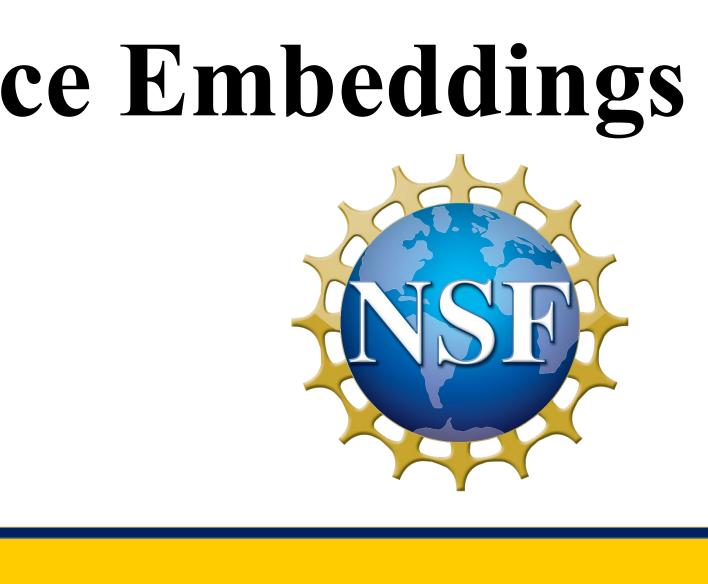
The gensim Python library was used to derive the 10 latent topic numbers for each turn The "AnotA" classifier identified (yes or no) whether dialogue act A was contained within the words of

Result

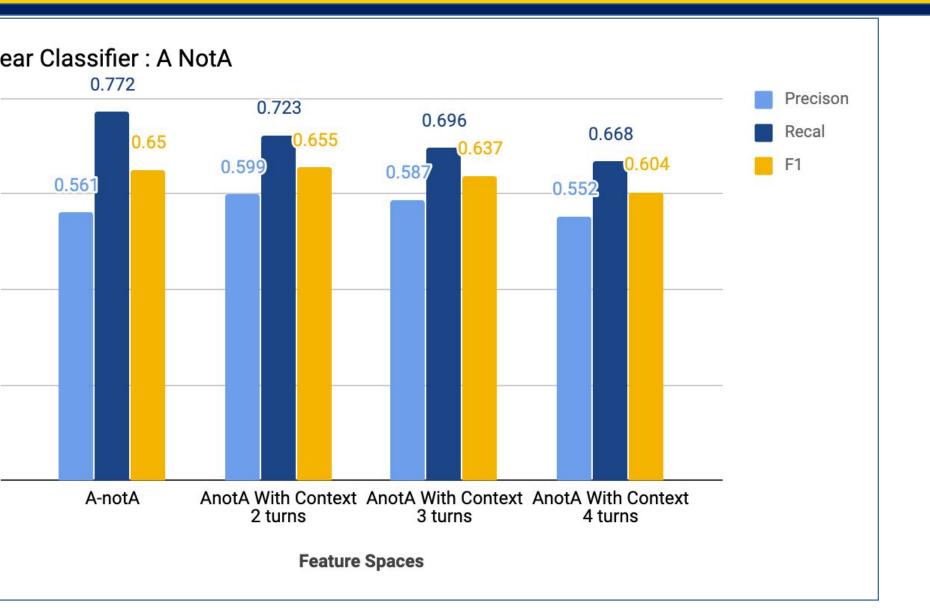
one dialogue turn.

- It achieved F1 accuracy (a combination of precision and recall) ranging from 0.6 for category A) sharing-ideas, to 0.3 for **D)** maintaining conversation.
- Adding 1 turn of context usually slightly improved results, adding more turns made it worse

| It ac cate 0.32 con | |
|---|---|
| | |
| Exp as c Tes Tun | |
| | |
| Parti by Impr (IUS Any reco are nece Scie Scie This grad | |
| Glas |) |



Doc2Vec Result



gure 5: Doc2Vec Results

ntence Embeddings Approach

o different sentences containing the same ords will appear differently to the classifier, king it more likely that the software will be le to recognize the different conversational tentions of the speakers.

esult

chieved F1 accuracy ranging from 0.65 for egory A) recognizing sharing-ideas, to for category D) maintaining

versation.

Future Work

pansion on using different classifiers such decision trees or random forest.

ting with different combination of features ning the Classifier to make an improvment

Acknowledgments

ial support for this work was provided the National Science Foundation's roving Undergraduate STEM Education E) program under Award No. 1504918. opinions, findings, and conclusions or ommendations expressed in this material those of the author(s) and do not essarily reflect the views of the National ence Foundation.

work was made possible thanks to luate student Duy Bui and Dr. Michael ss of Valparaiso University.