Looking for Low-Proficiency Sentences in ELL Writing

Shayne Miel
smiel@stanford.edu

Task

- Provide feedback to ELL writers on areas of low proficiency in their essays
- Previous work focuses on grammatical error detection/correction
- Not enough
  - Some sentences too jumbled to find specific errors
  - Some sentences sound unnatural even with no errors
- Hard to get right
  - CoNLL 2014 Shared Task (correcting grammar errors)
    - Best paper has F-score of 37.33 (Ng, et al, 2014)
    - Recent work has boosted to 47.40 (Rozovskaya and Roth, 2016)
    - Human agreement is extremely low, 45.91 - 72.58 F-score (Bryant and Ng, 2015)
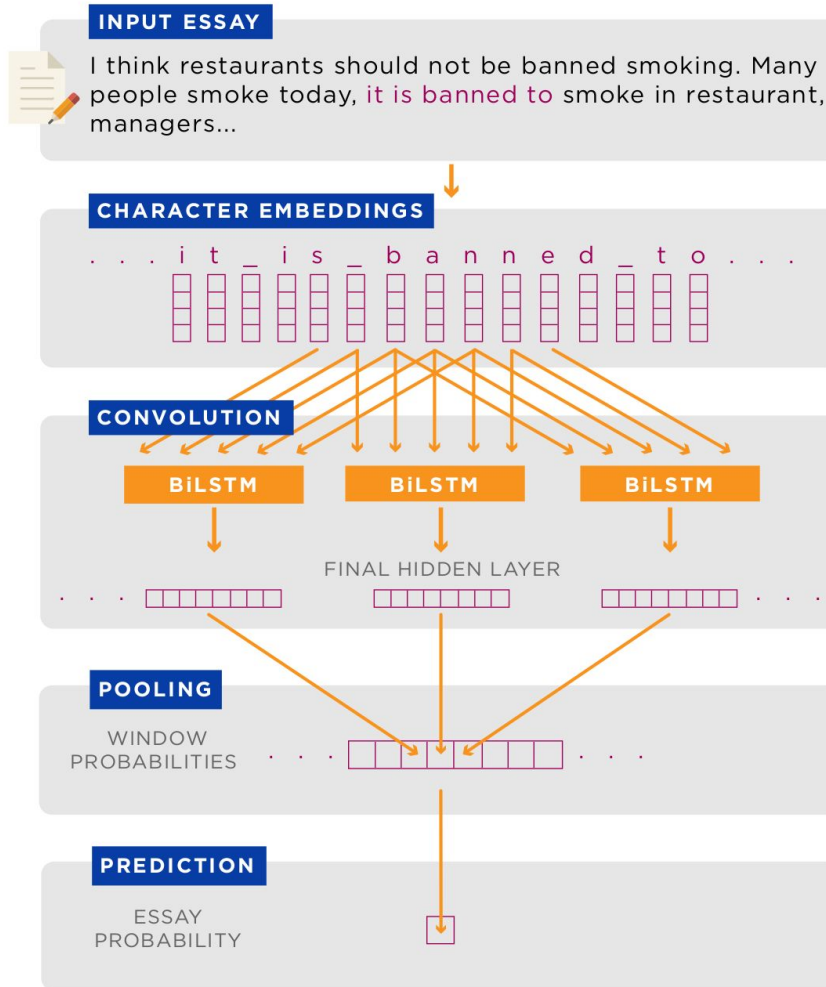
Approach

Trying an alternate approach

- What areas of the writing look the least like native English writing?
- Use a large dataset of native English and ELL student writing
- Train a convolutional network to predict whether the writing is ELL at the document level
- Hypothesis: I can use the window level predictions (output of the CNN *before* pooling) to highlight areas of the text that could use revision

Data

- Collected from 13 smaller data sets with wide array of L1s
- ICNALE, NUCLE, FCE, CEEAUS, MOECS, Gachon, TECCL, TOEFL-11, ASAP, and 4 private data sets from Turnitin.com
- Cleaning included:
  - Converting characters to ASCII and removing line breaks
  - Converting country names to "COUNTRY"
  - Removing excessively short or long essays
  - Discarding essays with keywords that were heavily biased towards one class
- Total of 77,046 essays after cleaning with 25 different L1s
- Test set: ICNALE, CEEAUS, MOECS   (6,133 essays, 3 prompts)
- Dev set: FCE, ASAP   (4,236 essays, 52 prompts)
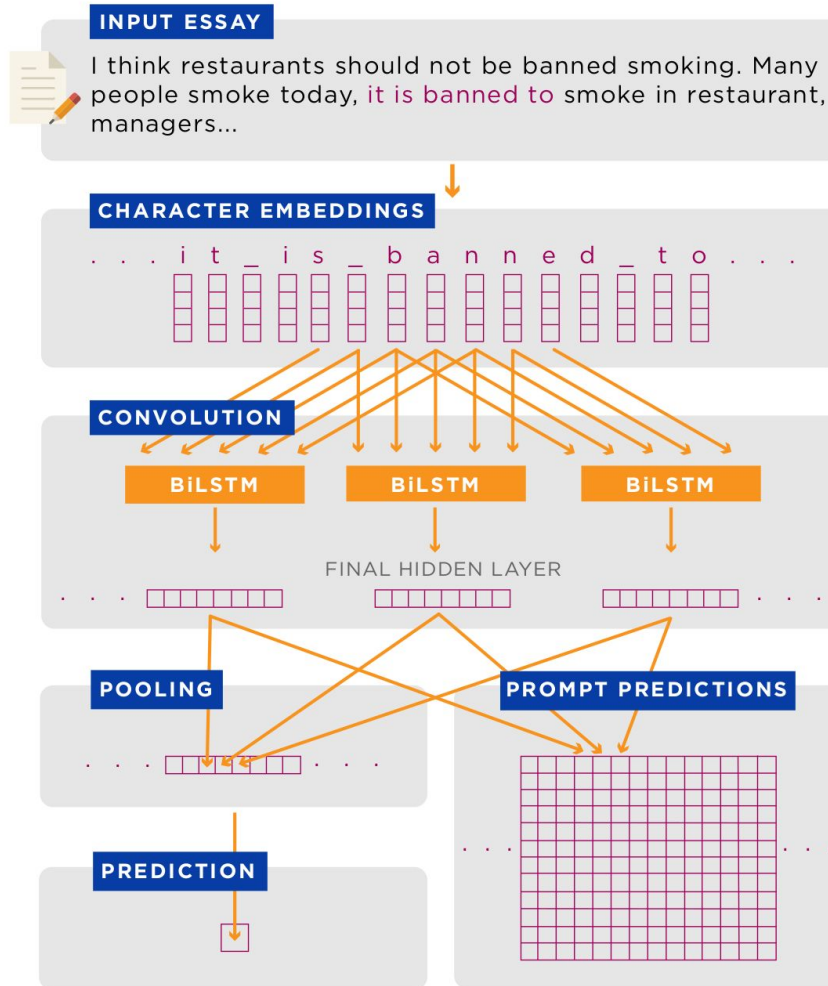- Train set: the rest   (66,677 essays, 95 prompts)

Models
Deep Filter

Models
Prompt Aware

INPUT ESSAY

I think restaurants should not be banned smoking. Many people smoke today, it is banned to smoke in restaurant, managers...

CHARACTER EMBEDDINGS

. . . i t _ i s _ b a n n e d _ t o . . .

CONVOLUTION

BiLSTM   BiLSTM   BiLSTM

FINAL HIDDEN LAYER

POOLING

PROMPT PREDICTIONS

PREDICTION

Models
Baseline

- Logistic Regression
  - Sort of
- Same CNN as described previously, except:
  - Input layer is character 4-grams per text window
  - Convolution filter is a single fully connected layer + sigmoid

Metric

- Area under the ROC curve
- At the document level (native English vs ELL)
  - On dev set to choose early stopping
  - On test set for model evaluation at the document level
- At the window level
- Hand annotated 200 randomly sampled windows from test set
  - Yes, give feedback = 1
  - No, don't give feedback = 0
  - AUC with window predictions from each model for model evaluation at the window level

Evaluation

- Model compared
  - Logistic Regression Baseline + max pooling
  - Logistic Regression Baseline + mean pooling
  - Deep Filter + max pooling
  - Deep Filter + mean pooling
  - Prompt Aware + max pooling
  - Prompt Aware + mean pooling
- Hyperparams consistent across models (details in paper)

Results

Model	Dev AUC	Test AUC	Window AUC
Logistic Max	0.711	0.569	0.583
Logistic Avg	0.771	0.796	0.586
Deep Filter Max	0.787	0.693	0.591
Deep Filter Avg	**0.922**	**0.861**	0.557
Prompt Aware Max	0.920	0.841	**0.607**
Prompt Aware Avg	0.917	0.856	0.597

Deep Filter Avg

Highest probability ELL essay   P(y=1) = 1.000

Donat need newspaper I think I donat need newspaper. To be honest, newspaper can keep long long time and it can read again and again if I donat throw away. But I think Internet is better than it. I have three reasons why I feel so. First, Internet can use sometime and somewhere. Recently, PC is light and thin, so it is very useful for having and go far away. Add to city have lots of wi-fi spot. For example, if I go to foreign country, I want to have Internet. It because I can get information about Japan in abroad. Second, I can get information quickly on the internet. Internet is one click then get information style. So it is so simple and easy way. And I can select best one of much information. For example, it is good for elderly people. They are very hard to see, but on PC letter size can change and easy to search. Younger people always use internet, too. So, it is very convenient for everyone. Third, Internet must not miss information. On the internet, it have lots of information. It because, information must not disappear. So, nowadays information and news can keep without newspaper and magazin. For example, I can find news which is 10 years ago and so on. But on the other hand, I have to use internet very carefully. Information which is on the internet stay forever. I think both of them have good side and bad side. In constance, newspaper can keep long time. But on the other side it is hard to get information and news. Internet is very useful and smart. By the way, I have to have a lot of attention to using it. But I think internet have more good points than newspaper. For those reasons, I think I donat need newspaper.

Prompt Aware Max

High probability windows

part time job can be continued as students' carrier when students are not sure about their future an
some money by ourselves work. We use the money important which we get by our work. We understand tha
ney given by my parents and I didn't know it is hard to make money. After becoming 20, I work hard t
portant for us definitely. For we have working experience rather than not experienced. And, experien
advantages. In the first place, you can practice your ability of communication. When you work, you w
aurants, it is absolutely a place for non toxic areas, thus, the air should be fresh and smell clean
 how hard earn money and live in the future. By those reason, I think that having a part-time job is
y, having a part-time job sometimes damages students' health. As I mentioned above, because of their
t be necessary for the older generations who are not tech-savvy. However, skills for using the Inter
and got a job. What we can have a part-time job during a college student is a very good chance. Both
t all the restaurants in the country. In my opinion, there are several reasons to support my answer.
ly on it for communication and accomplishing our daily tasks. Although it seems like the Internet ha
 et some job? My answer was Yes, I have.. And I really think that everyone used to think. May be now
y can go to college. Actually, I started to live alone near my college and to do a part-time job. It
 ot good for person's body. And smoking is weakened breathing smoke fitness comes the hard mucus. In
 hinder one's studying process. Of course, this is not to completely ignore the benefits of having a
comparable to the effects of food poisoning, if not more severe. To conclude, the long-term side eff
provide customers with a nice environment to consume their meals, smoking should not be allowed as i

Prompt Aware Max

High probability windows (transitional phrases)

part time job can be continued as students' carrier when students are not sure about their future an some money by ourselves work. We use the money important which we get by our work. We understand tha ney given by my parents and I didn't know it is hard to make money. After becoming 20, I work hard t portant for us definitely. For we have working experience rather than not experienced. And, experien advantages. In the first place, you can practice your ability of communication. When you work, you w aurants, it is absolutely a place for non toxic areas, thus, the air should be fresh and smell clean how hard earn money and live in the future. By those reason, I think that having a part-time job is y, having a part-time job sometimes damages students' health. As I mentioned above, because of their t be necessary for the older generations who are not tech-savvy. However, skills for using the Inter and got a job. What we can have a part-time job during a college student is a very good chance. Both t all the restaurants in the country. In my opinion, there are several reasons to support my answer. ly on it for communication and accomplishing our daily tasks. Although it seems like the Internet ha et some job? My answer was Yes, I have.. And I really think that everyone used to think. May be now y can go to college. Actually, I started to live alone near my college and to do a part-time job. It ot good for person's body. And smoking is weakened breathing smoke fitness comes the hard mucus. In hinder one's studying process. Of course, this is not to completely ignore the benefits of having a comparable to the effects of food poisoning, if not more severe. To conclude, the long-term side eff provide customers with a nice environment to consume their meals, smoking should not be allowed as i

Prompt Aware Max

High probability windows (transitional phrases)

part time job can be continued as students' carrier when students are not sure about their future an
some money by ourselves work. We use the money important which we get by our work. We understand tha
ney given by my parents and I didn't know it is hard to make money. After becoming 20, I work hard t
portant for us definitely. For we have working experience rather than not experienced. And, experien
advantages. In the first place, you can practice your ability of communication. When you work, you w
aurants, it is absolutely a place for non toxic areas, thus, the air should be fresh and smell clean
how hard earn money and live in the future. By those reason, I think that having a part-time job is
y, having a part-time job sometimes damages students' health. As I mentioned above, because of their
t be necessary for the older generations who are not tech-savvy. However, skills for using the Inter
and got a job. What we can have a part-time job during a college student is a very good chance. Both
t all the restaurants in the country. In my opinion, there are several reasons to support my answer.
ly on it for communication and accomplishing our daily tasks. Although it seems like the Internet ha
et some job? My answer was Yes, I have.. And I really think that everyone used to think. May be now
y can go to college. Actually, I started to live alone near my college and to do a part-time job. It
ot good for person's body. And smoking is weakened breathing smoke fitness comes the hard mucus. In
 hinder one's studying process. Of course, this is not to completely ignore the benefits of having a
comparable to the effects of food poisoning, if not more severe. To conclude, the long-term side eff
provide customers with a nice environment to consume their meals, smoking should not be allowed as i

Conclusion

- Good news
  - The Deep Filter network does a significantly better job than the baseline at distinguishing between ELL and native English essays.
  - At the essay level, the network does seem to be picking up on extremely low-proficiency essays.
- Bad news
  - Window level AUC barely better than guessing
  - My hypothesis that the window predictions from this classifier would enable me to highlight specific areas of low proficiency did not work out
- Why?
  - ELL students who are struggling with putting their ideas down in English are more likely to lean on the crutch of standard transitional phrases. These cues are easier for the model to pick up on than the varied ways a student might introduce errors into a sentence.

Next Steps

- Try using a language model trained on well-edited English text to see if areas of high perplexity correlate with areas of low proficiency in ELL writing.
- Collect more sentence level proficiency annotations so that I can treat this as a fully supervised process, rather than using the document's ELL status as a proxy.
- Implement the methods in "*Predicting Grammaticality on an Ordinal Scale*" (Heilman, Michael, et al., 2014) to see if areas of low proficiency are related to the amount of difficulty a reader would have understanding the sentence due to the grammatical errors.

Thank you!

Shayne Miel

smiel@stanford.edu