

Formality Style Transfer with Deep Learning

Yonglin Wang



Too formal...

~~Hello, world!~~

Hey, world, you know!

What's formality anyways :o

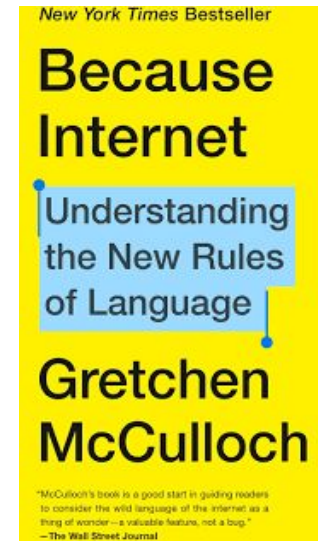
Scope of this project: specifically, informal speech on the internet

Intergenerational communication

- Internet exposure affects the way we talk
- Interpreting and bridging the gap

English as a Second Language

- A lot of nuance to navigate, not taught in mainstream textbook
- "I shall be writing like such in my GRE tests cuz I sincerely believe that it's super duper native yet elegantly sophisticated."
- Instructional tool for learning formality



GYAFC (Grammarly's Yahoo Answers Formality Corpus)

- Total of 110K informal-formal sentence pairs collected from Yahoo! Answers (as of 10/25/2007!)
- Hire Turkers to provide rewrites

Informal	Formal
He's supposed to be in jail!	He is supposed to be in jail.
I dun think he loves her	I do not think he loves her.
The sound of the 2 together, wow I do love the sound lol =)	I love the sound of the two together.
I love Tyra, she has a great show, but Oprah is just great!	I love Tyra's show, but I prefer Oprah.
How dare he call her kids - orphans.	How dare he refer to her children as orphans?

Tasks for the System

End goal: build a pipeline that....

- Classifies the input sentence
 - [FastText Classifier](#)
 - Trains literally in *seconds*, on CPU!
 - turned out to be good enough: 0.88 accuracy, ~8 MB model file after quantization
- Suggests rewrites in the other style
 - Transformer vs. a VAE style-transfer model
 - Much more difficulties

Why is rewrite hard?

In theory...

How many informal sentences can be mapped to a formal one?

- “There are simply so many.”
- “There’s just so many.”
- “There’s jsut soo mnay.”
- “There’s just SOOOO MANY!!”
- “there’s just, i mean, definately SOOOO MANY!!”

Well, isn’t that what the fancy deep learning models are here for?

Approach 1: Casting the task as Machine Translation

Why a language translation task?

- Language translation has this “soooo many” to one mapping too
- No complicated engineering required; it’s just fairseq >:D

Use a transformer model--but smaller!

- `transformer_iwslt_de_en`
- A smaller FFN dim: 2048 (original) -> 1024 (ours)

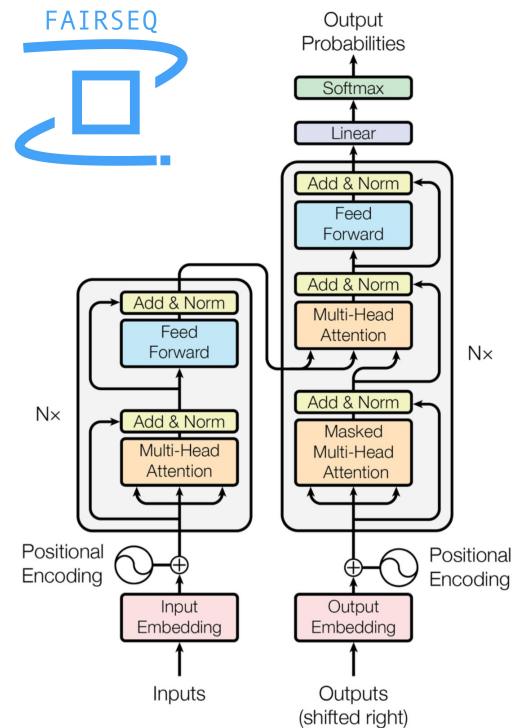
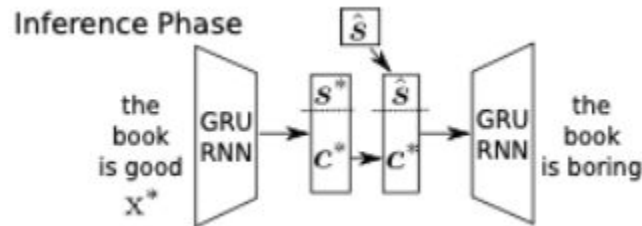


Figure 1: The Transformer - model architecture.

Approach 2: Casting the task as Style Transfer

- Still “Deep”, but more delicate and specialized
- Uses VAE and RNN
- Input \rightarrow [style; content] \rightarrow Rewrite
- For sentiment transfer, it seems to work



Evaluations

	Transformer (transformer_iwslt_de_en)			Disentangled-VAE		GYAFC Best Baseline (LSTM)
	BLEU	BLEU (lc)	Ratio, gen/ref	BLEU (lc)	Ratio, gen/ref	BLEU (case-NA)
Informal -> Formal	68.9	71.1	0.998	6.7	0.834	67.67*
Formal -> Informal	37.5	43.0	1.039	4.8	0.952	NA

*Evaluated based on a subset of test set. Casing & tokenization not specified.

Test Set Output: Informal -> Formal

Input (informal)	Tran. Output (formal)	VAE Output (formal)	Reference (formal)
I LOOOOOVVVVVVVEEE this song SOOO Much!!!!!!	I am very fond of this song.	i love the song	I love the song so much.
I dun think he loves her	I do not think he loves her.	i believe that he is attractive	I do not think he loves her.
The sound of the 2 together, wow I do love the sound lol =)	I love the sound of the two together.	the sound of the sound of the sound of the sound of the sound	I love the sound of the two together.
I love Tyra, she has a great show, but Oprah is just great!	I love Tyra because she has a great show, but Oprah is great!	i like oprah oprah oprah oprah oprah is oprah	I love Tyra's show, but I prefer Oprah.
How dare he call her kids - orphans.	How dare he call her kids - orphans.	how can i eat children	How dare he refer to her children as orphans?

Test Set Output: Formal -> Informal

Input (formal)	Tran. Output (informal)	VAE Output (informal)	Reference (informal)
I am still fond of both bands today.	I still like both bands today.	i like the bands today	still like em both!
I suggest you let him know that you want to marry him.	tell him you want to marry him.	i don't know if you want to marry him	Tell him you wanna marry him.
I love all of the songs of Green Day, and particularly Boulevard of Broken Dreams.	I love all the songs of Green Day and especially Boulevard of Broken Dreams.	i love the song i love the song i love the song i love	I love all Green Day songs especially Blvd of broken dreams
Not I; however, I was previously in orchestra.	not me but i was in orchestra.	no i like the olsen twins	Not me, I used to be in orchestra though.

Assembling a Two-Step Pipeline

- Input -> fasttext formality classifier -> transformer rewrite
- Wished for a Pythonic API experience

Assembling a Two-Step Pipeline

- Input -> fasttext formality classifier -> transformer rewrite
- Wished for a Pythonic API experience

My pipeline, with wrappers for docker server and subprocesses:



Demo Time!

TEXTING:
heyyyyyy

FRIENDS I THINK I LIKE YOU TAKE THE HINT, LET'S PLAY I'M DRUNK DUMBASS\$

heyyyyyy

Hello. Hyy is a very good singer. He is too nice. He is not very nice. I am way too much in favor of that.

Credit: upper: online; lower: Xiaoyu

Next Steps...

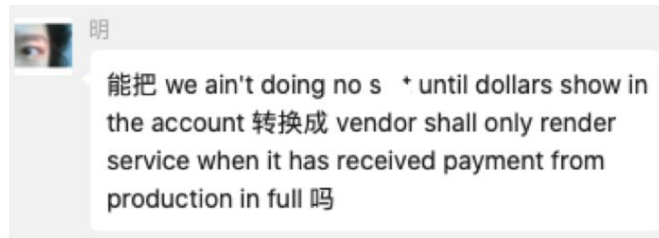
- Apply to passages (i.e. groups of sentences)
 - Sentence tokenization is _____
 - Especially with... (drum roll plz!!) informal texts!!!
- Generalize to other domains
 - Currently: Internet Language (Family and Relationship, Entertainment and Music)
 - Possible directions: Emails, Academic writing, legal documents

DECIPHERING ACADEMESE

YES, ACADEMIC LANGUAGE CAN BE OBTUSE, ABSTRUSE AND DOWNRIGHT DAEDAL. FOR YOUR CONVENIENCE, WE PRESENT A SHORT THESAURUS OF COMMON ACADEMIC PHRASES

"To the best of the author's knowledge..."	=	"WE WERE TOO LAZY TO DO A REAL LITE-RATURE SEARCH."	"It should be noted that..."	=	"OK, SO MY EXPERIMENTS WERENT PERFECT. ARE YOU HAPPY NOW??"
"Results were found through direct experimentation."	=	"WE PLAYED AROUND WITH IT UNTIL IT WORKED."	"These results suggest that..."	=	"IF WE TAKE A HUGE LEAP IN REASONING, WE CAN GET MORE MILEAGE OUT OF OUR DATA..."
"The data agreed quite well with the predicted model."	=	"IF YOU TURN THE PAGE UPSIDE DOWN AND SQUINT, IT DOESN'T LOOK TOO DIFFERENT."	"Future work will focus on..."	=	"YES, WE KNOW THERE IS A BIG FLAW, BUT WE PROMISE WE'LL GET TO IT SOMEDAY."
			"...remains an open question."	=	"WE HAVE NO CLUE EITHER."

JORGE CHAM © 2004
www.phdcomics.com



Lessons Learned

- If one fancy deep learning model doesn't work...
 - Don't just tune the hyperparameter, especially if the baseline is abysmal
 - Try approaching from a different angle, e.g. recasting it as a different task
- Your dataset matters!
 - Legit corpus; sometimes biased and outdated very quickly
- Remember to ask for help!
 - "Try for x minutes and then reach out"
 - Without Prof. Lignos' generous instructions, I could have been...
 - Wrestling with a DistilBERT classifier on CPU
 - Implementing Transformers from scratch w/ PyTorch
 - Smashing my keyword because fastText and Fairseq refused to cooperate

References

John, V., Mou, L., Bahuleyan, H., & Vechtomova, O. (2018). Disentangled representation learning for non-parallel text style transfer. *arXiv preprint arXiv:1808.04339*.

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Rao, S., & Tetreault, J. R. (2018, January). Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer. In *NAACL-HLT*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *arXiv preprint arXiv:1706.03762*.

Thanks for Tuning in :)



The model thinks you are 100% **informal** .