

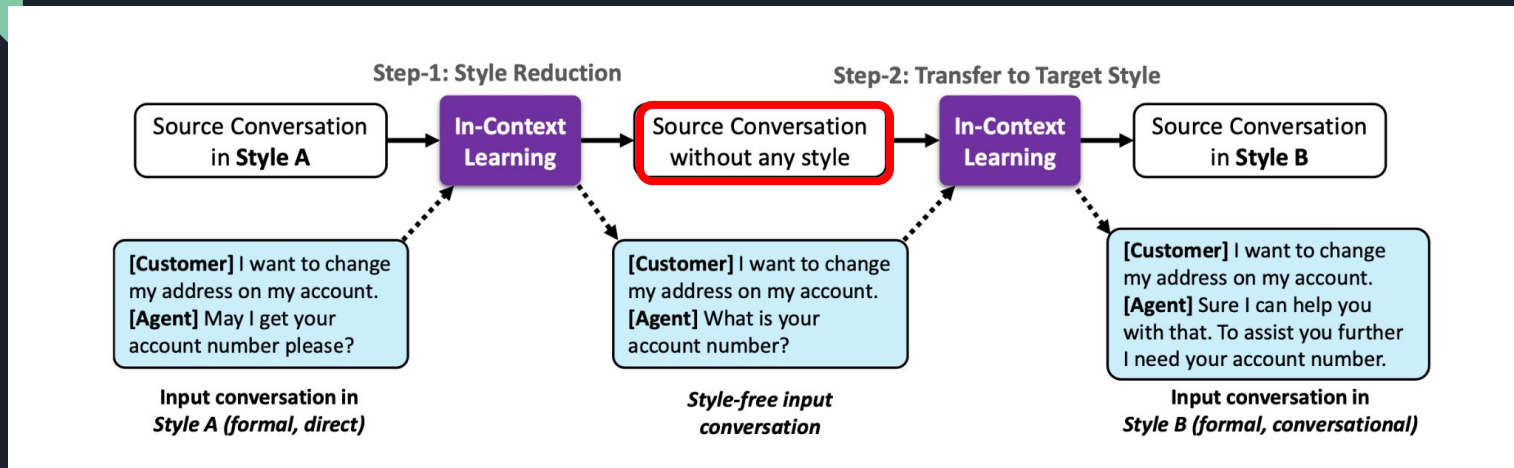
Automating Style-Free Text Generation



Cameron Yenche
DATASCI 266

Berkeley
SCHOOL OF
INFORMATION

AWS Achieves Few-Shot Style Transfer



- AWS AI Labs generated a model for few-shot style transfer
- Generating Style Free Text using LLMs required human verification



Purpose for our Project:

Build a more **Robust Automation of
Style-Free Text Generation** to allow for
end-to-end few-shot style transfer in
Commercial Chatbots



Kaggle @Kaggle

Customer Support Twitter Dataset:

- Contains Support Tweets from 20 Top Brands
- Provides Representative text that commercial ChatBots would find during deployment

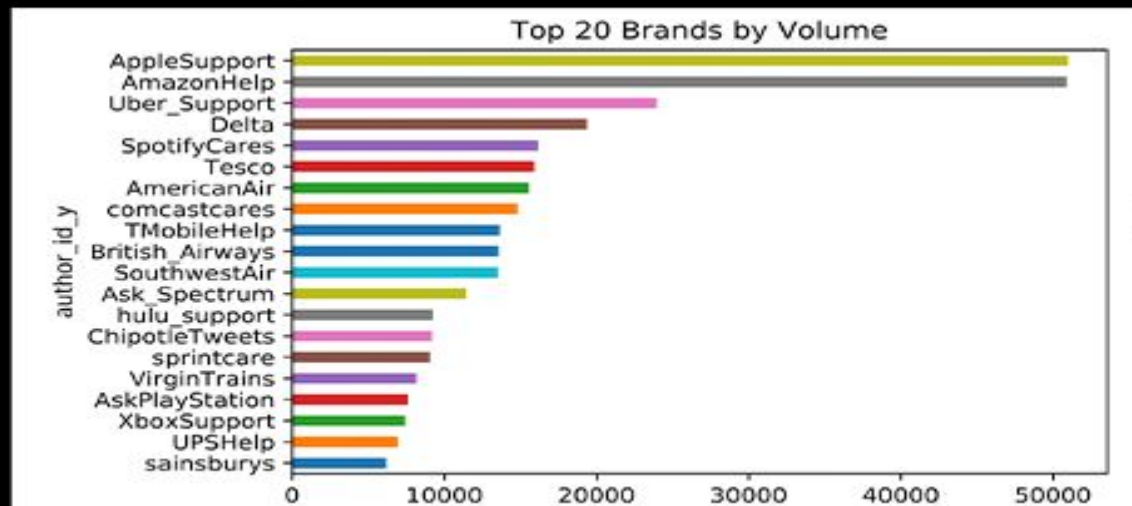
Key Data Fields:

- tweet_id
- text ... [Show more](#)



Kaggle @Kaggle

Plot of Represented Brands:





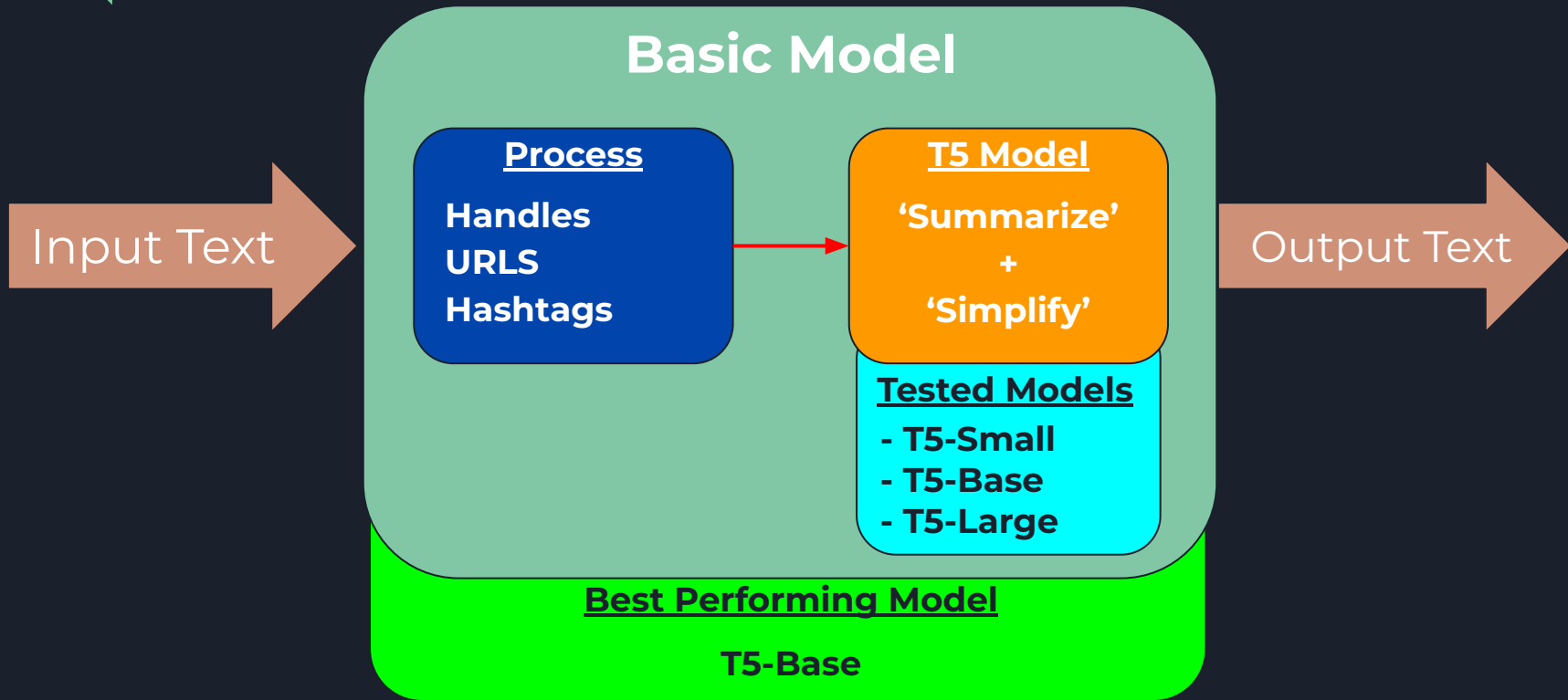
Model Performance and Style Metrics

- **BERT Score:** assesses how well meaning is transferred.
- **Flesch Reading Score:** assesses complexity and readability
- **Lexical Diversity:** measures prose diversity, serves as a proxy for style

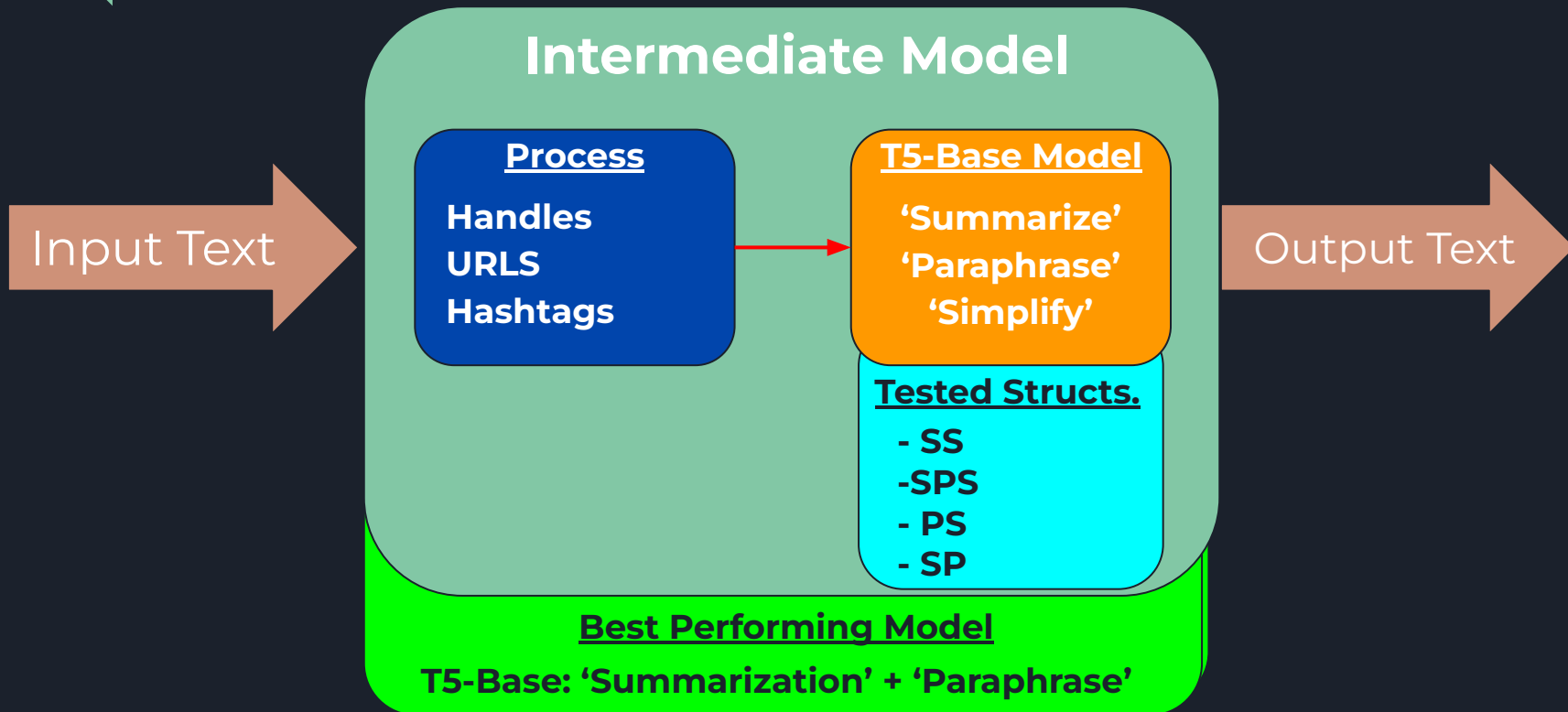
GOAL:

- Maintain  Increase  Reduce 

Basic Model Structure



Intermediate Model Structure

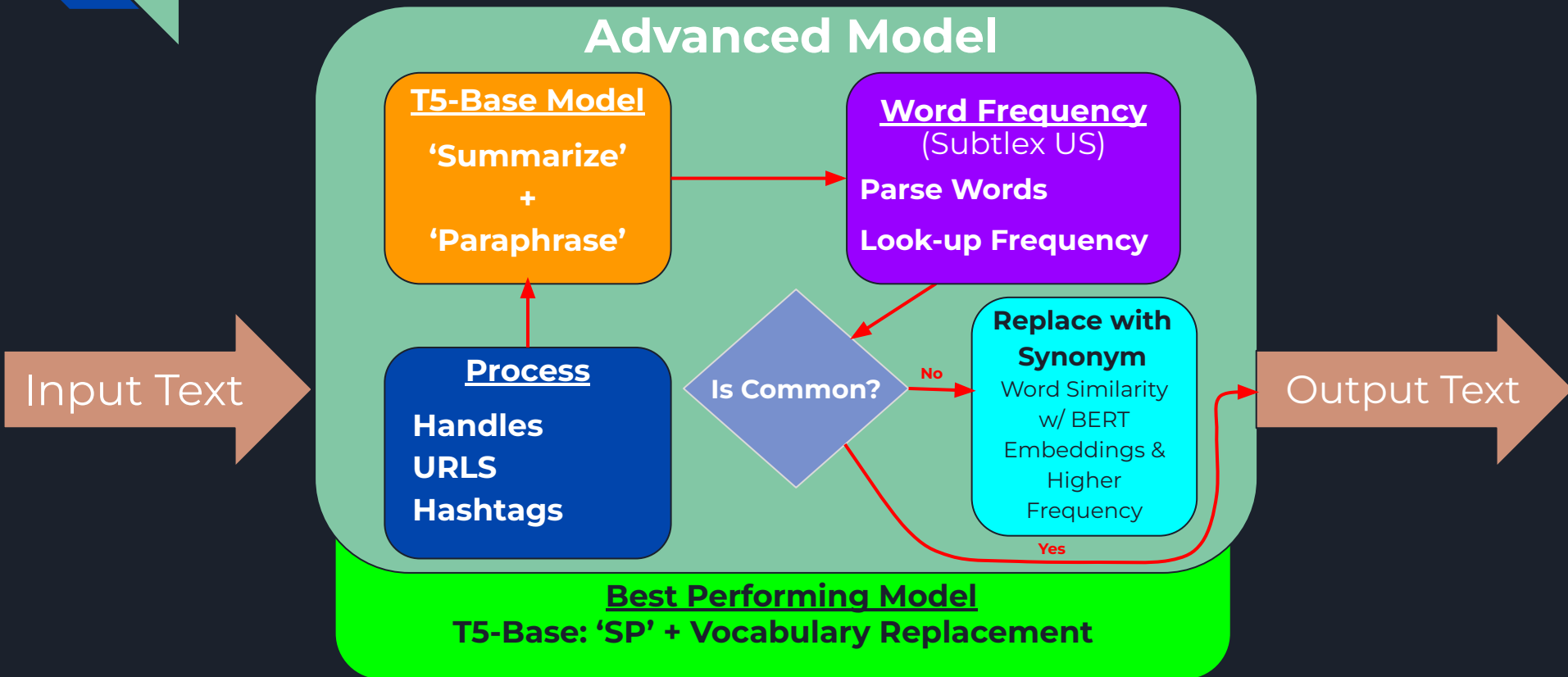



Subtlex US Dataset

- Corpus of 51 million movie lines
- Representative of the English Language for word frequency dictionary



Advanced Model Structure





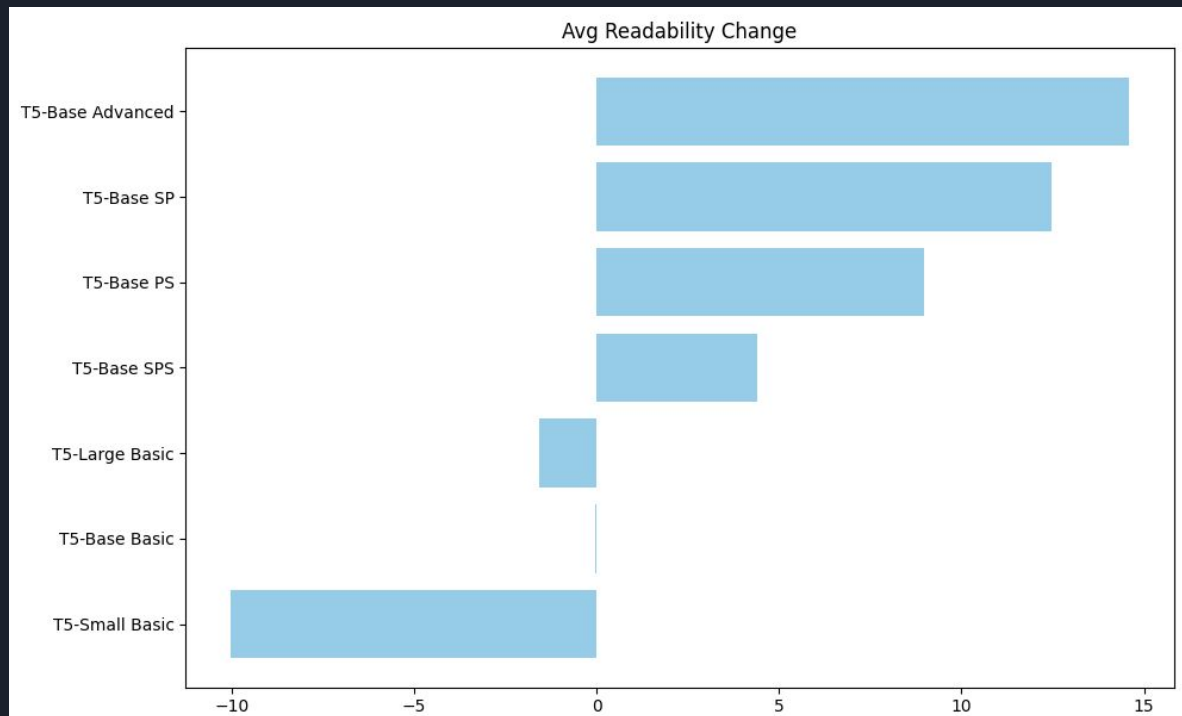
Amidst the **ethereal twilight**, the **celestial** bodies
danced in harmony, their radiant hues painting the
canvas of the sky with breathtaking elegance, while
the gentle zephyrs whispered secrets of ancient lore.



the **heavenly** body dance in harmony
in the heavenly **fall** .

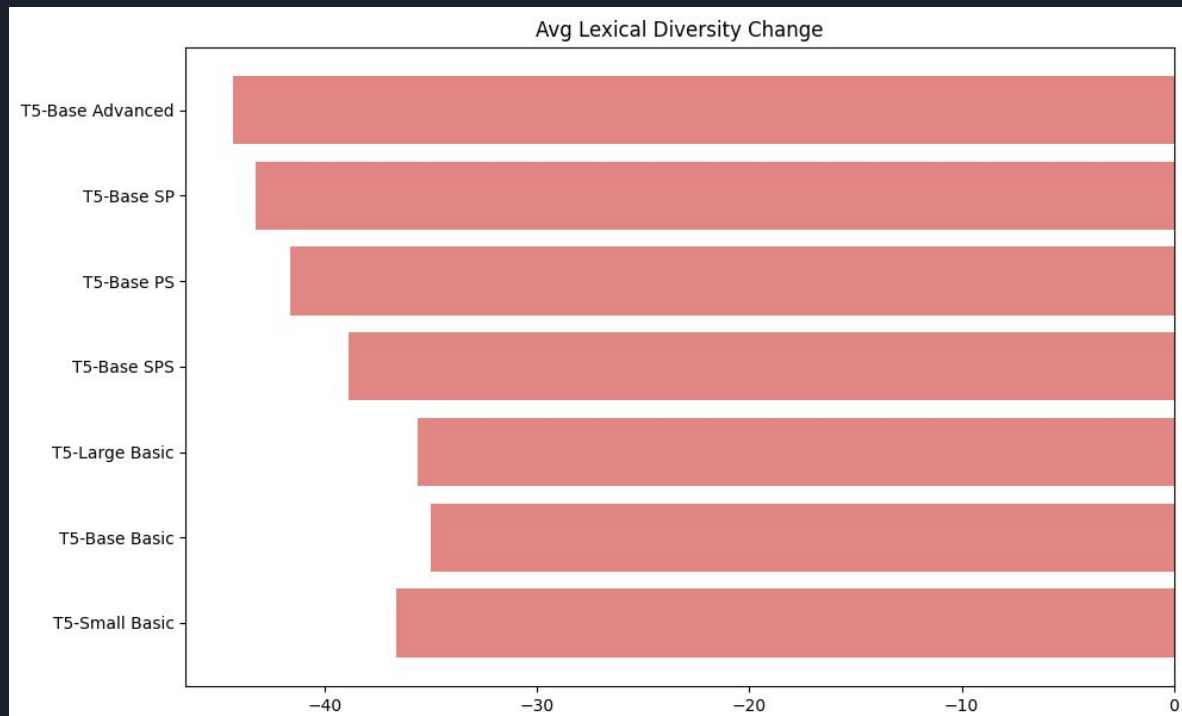
Metrics Evaluation Across Validation Set

Model Complexity



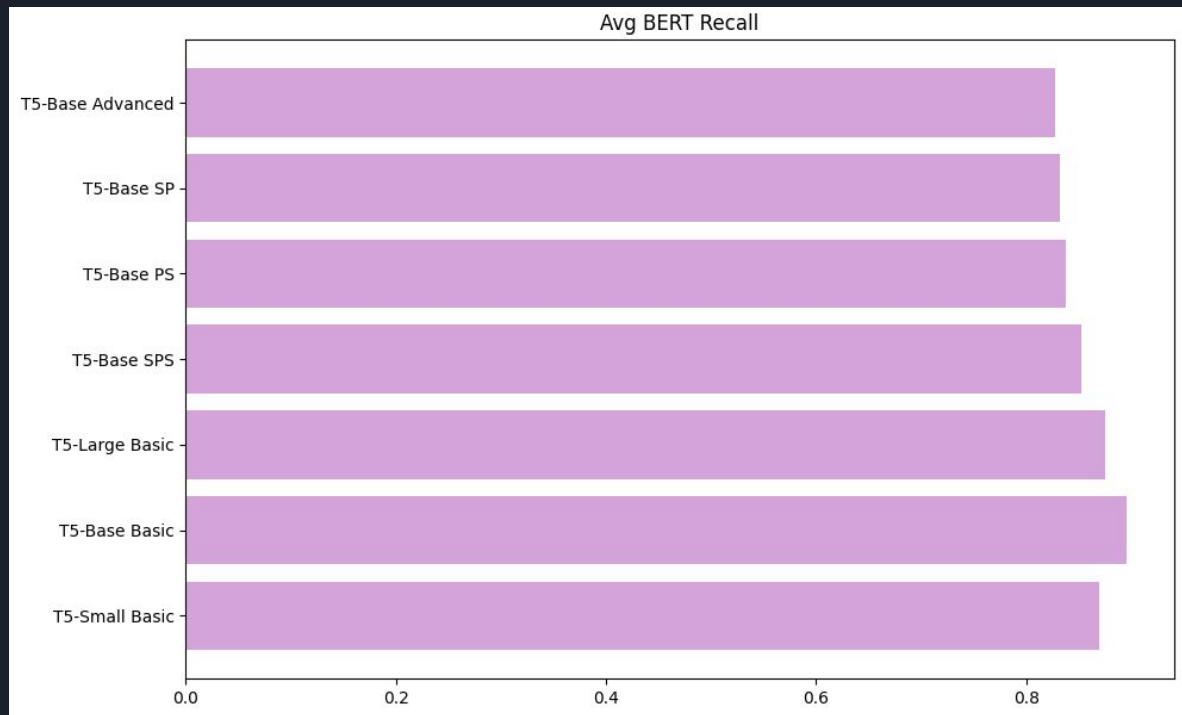
Metrics Evaluation Across Validation Set

Model Complexity



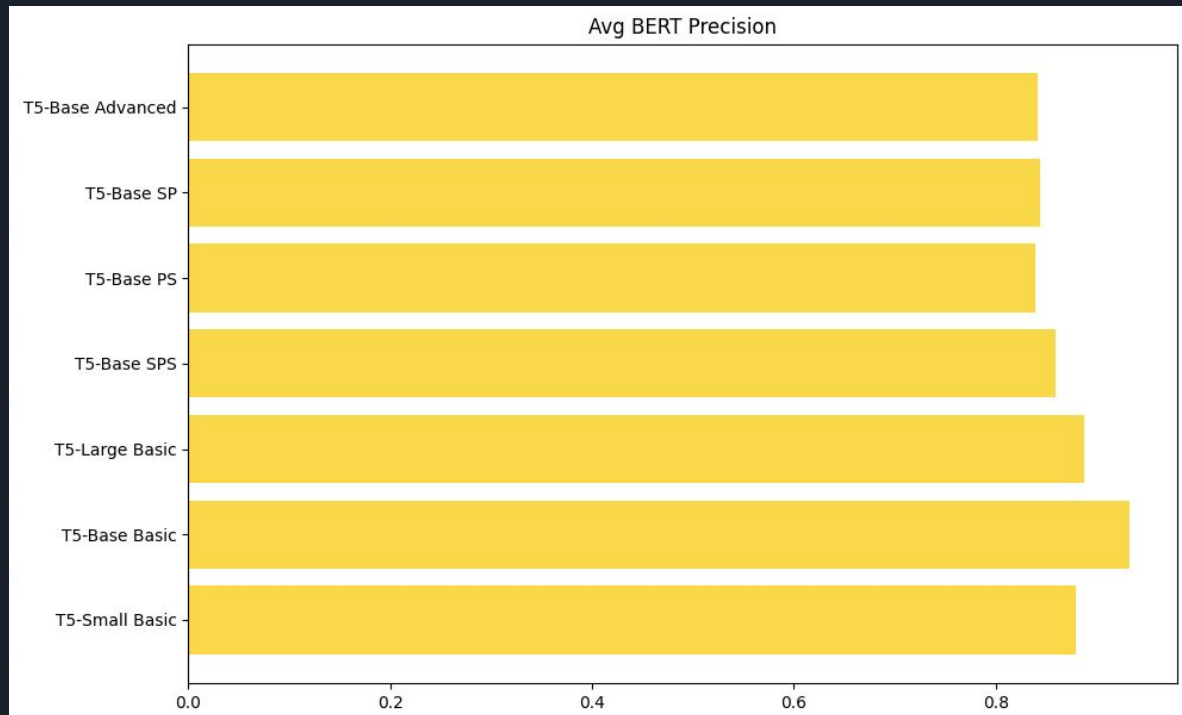
Metrics Evaluation Across Validation Set

Model Complexity

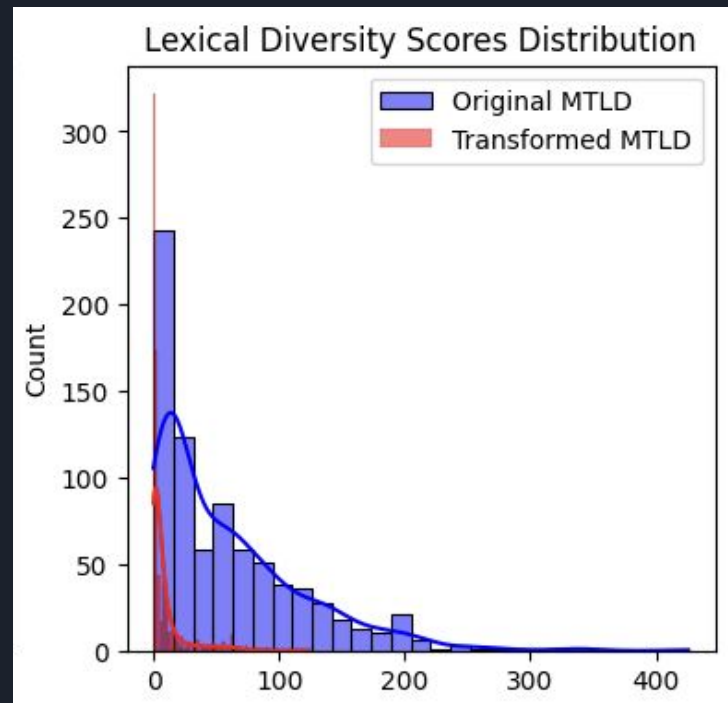
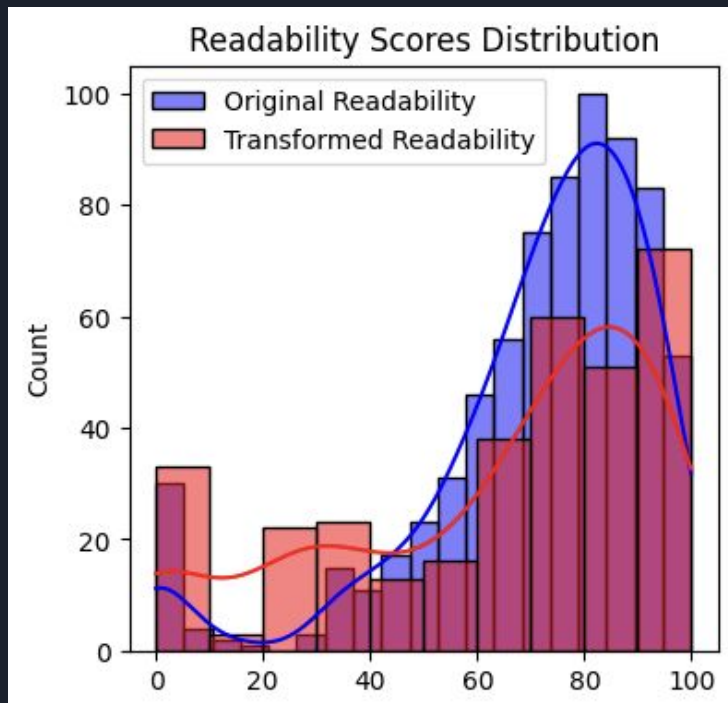


Metrics Evaluation Across Validation Set

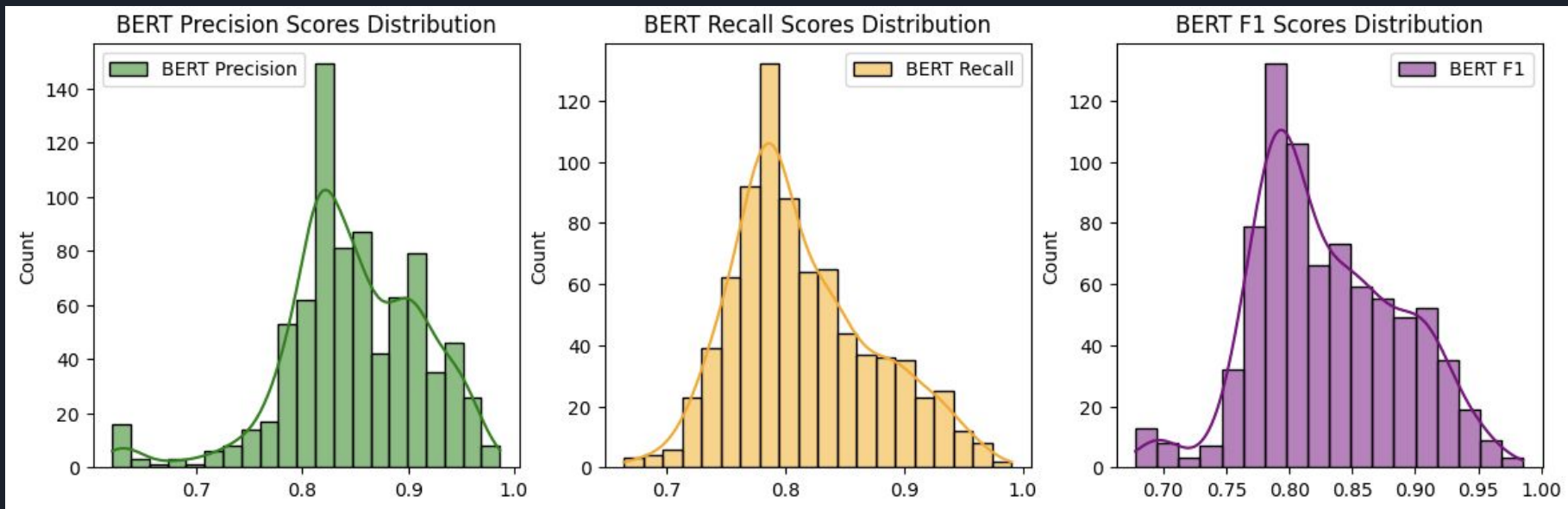
Model Complexity



Style Metrics on Test Set



BERT Scores on Test Set





Thank You!

Berkeley SCHOOL OF
INFORMATION



Appendix

1. Next Steps & Future Work
2. (For Fun) Full Photoshopped X (Twitter) Kaggle Slide



Next Steps & Future Work

- Improve Runtime (4-7 seconds is unacceptable an intermediate conversion step)
- Fine tune a T5 (was explored during this project, computationally costly)
- Expand the model s to incorporate emojis as well as multi-modal, multilingual input



Home



Explore



Notifications



Messages



Lists



Bookmarks



Top Articles



Communities



Verified Orgs



Profile



More

Tweet



Kaggle
@Kaggle



Kaggle
@Kaggle

What's happening?!

everywhere about.twitter.com Born March 21
Joined February 2007

0 Following 65.7M Followers

Followed by Raws, AI Authority, and 461 others you follow

Tweets

Affiliates

Replies

Highlights

Media

Likes



Kaggle @Kaggle

Customer Support Twitter Dataset:
- Contains Support Tweets from 20 Top Brands
- Provides Representative text that commercial ChatBots would find during deployment

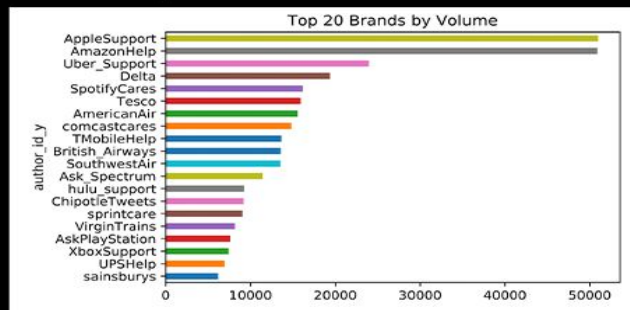
Key Data Fields:

- tweet_id
- text ... [Show more](#)

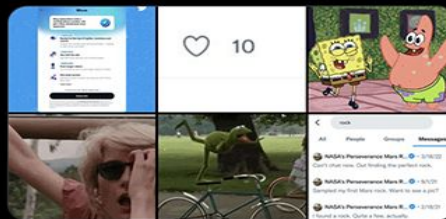


Kaggle @Kaggle

Plot of Represented Brands:



Search Twitter



Listen live in Spaces

#FinanceDaily: China troubles | Twitter is dead, long live X | Fed Hike

+4.6K

Coffee with Captain from @dGenNetwork

+165

You might like



Bill Gates
@BillGates

Follow



Android
@Android

Follow



Twitter Support
@TwitterSupport

Follow

Show more

What's happening

FIFA Women's World Cup · LIVE
Brazil vs Panama



Messages

Trending in United States