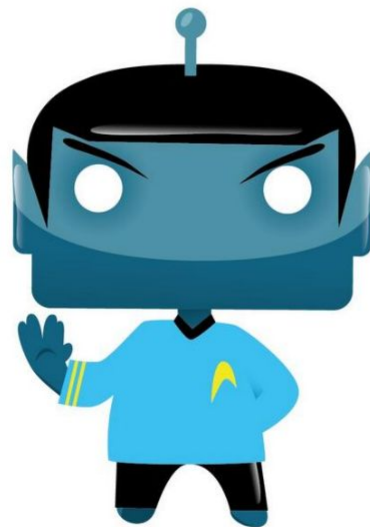


# Enterprise to Computer : Star Trek Chatbot



# Introduction

- Chatbots used in multiple domains.
- Adding persona to make chatbots more human like.
  - Counsellor -> ?
  - TV show/movie chatbots -> ?
- Propose **E2Cbot**:
  - Captures Star Trek characters' linguistic style.
  - Incorporates references from the show.
  - Data driven.



# Agenda

- Attempt to capture characteristic tones of Star Trek characters.
- Data driven vs Rule based chatbot.
- Try to mix both.

# Pandora Bot

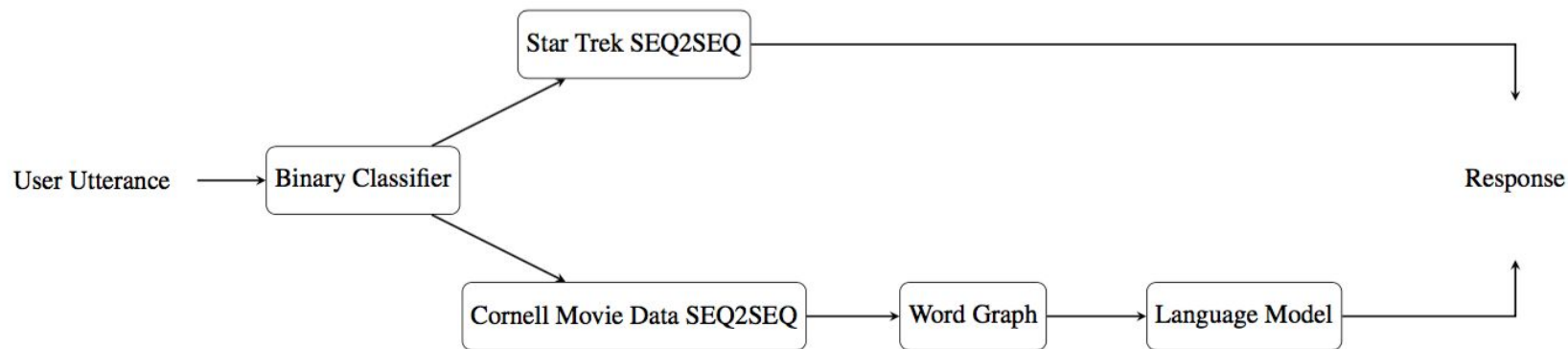
- Fake Spock bot imitates Spock.
- Rule-based.
- Too complex to create and update the rules.
- Might not cover all cases.

User	Fake Spock Bot
I am hungry captain.	What will you be eating?
You can do it.	I shall attempt it.
Mister Spock.	I am an AI chatting on the Internet.
Maximum power.	I see no logic in preferring Stonn over me.

Table 1: Responses by baseline Pandora bot.



# Design Schematic



# Dataset

- The Star Trek dialog dataset.
  - Created our own dataset, pulled from various Star Trek T.V. episodes and movies scripts
  - Same exchange between characters gives multiple pairs
  - Exchange  $A \rightarrow B \rightarrow C$  gives (A, B), (B, C) and (AB, C) post reply pairs.
  - 100,990 post-response pair.
  - 14.3 average words per utterance.
- Cornell Movie-Dialogue Corpus.
  - Conversational exchanges between movie characters.
  - 199,455 post-response pairs.
  - Average utterance length 12.82 words.

# Dataset

- Open source Twitter dataset.
  - Used only to train the binary classifier.
  - For current topics not there in Cornell Movie dataset
  - 50,000 post response tweet pairs.
  - Average utterance length 16.18 words.

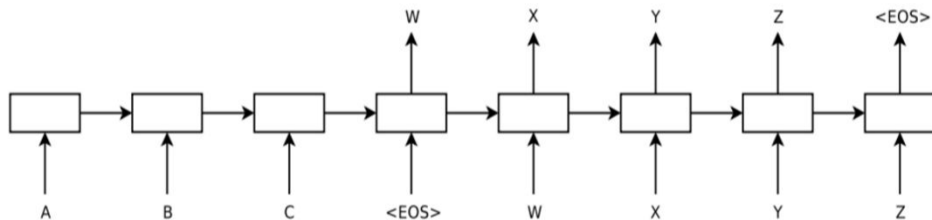
# Binary Classifier

- Logistic regression-based
- Routes utterance to Star Trek SEQ2SEQ or Cornell Movie Data SEQ2SEQ
- Training Data
  - 200,000 Star Trek dialogs
  - 100,000 Cornell Movie Dialog Corpus dialogs
  - 100,000 tweets from the Twitter dataset
  - Randomly sampled 80% dialogs as training and 20% dialogs as test data.
- Feature space:
  - Top 10,000 TF-IDF unigrams and bigrams.
  - Removed stop words.
- 95% accuracy on the test set



# SEQ2SEQ Models

- For Star Trek style input:
  - Trained on Star Trek dialogs dataset.
  - 17400 global steps.
  - Perplexity 18.21.
- For everyday conversations:
  - Trained on Cornell Movie dialogs dataset.
  - 27800 global steps.
  - Perplexity 12.48.



Hardware: AWS EC2 pc2.xlarge instance with one GPU which is a NVIDIA Tesla K80 Accelerator.

# Word Graph

- Style shifting normal conversation domain to Star Trek domain.
- Constructed using Star Trek dialogs.
  - Stores words and their POS tags as nodes
  - Adjacency as edges
- Looks up word graph to insert appropriate and frequent words between adjacent words in the input.
- Good at inserting names at start and end.

I am sorry **Miranda**  
I will go **back**  
**Uhura** how are you  
**Captain** shall I leave

Figure 2: Generated responses for normal conversation SEQ2SEQ. The words in bold have been added by the Word Graph algorithm.

# Filtering Word Graph Responses

- Word Graph constructs many ungrammatical and incoherent sentences.
- Language Model:
  - Sentences are ranked using a Bigram Language Model trained on Star Trek dialogs to find the most likely response.
  - Language model filters out impossible grammatical structures.
  - If multiple sentences have highest probability, choose the one containing words present in a handcrafted keyword list.
- Perplexity:
  - If response perplexity is very low or very high when compared to perplexity of Star Trek dialogs ( $>50$  or  $<10$ ), select a response from standard response set.

# Evaluation

- Quantitative metrics:
  - Perplexity
  - Overlap with Star Trek vocabulary
- Human evaluation:
  - We used a set of 20 standard input sentences:
    - 10 regular everyday sentences.
    - 10 star trek like sentences.
  - Human annotators rate the responses on basis of:
    - Correct grammar.
    - Coherence or relevance.
    - Star Trek relatedness.

# Evaluation

- Score of 0 if the response does not exhibit the property or 1 if it does.
- Six Star Trek fans, four not familiar with Star Trek.
- Used the Pandora Bot as a baseline comparison.

# Results

Model	Average Perplexity	Vocabulary Overlap
Pandora Bot	45.77	95.31%
E2Cbot	60.92	91.45%

Table 3: Average perplexity and vocabulary overlap of various models on the Star Trek dataset.

- Perplexity of Star Trek dialogs was 65.69.

Metric	Pandora Bot	E2Cbot
Grammar	<b>98%</b>	93.5%
Coherence	60.5%	<b>73.5%</b>
Star Trek style	64%	<b>86%</b>
Average score	74.16%	<b>84.33%</b>

Table 2: The human evaluations to compare the Pandora bot with the E2Cbot model.

	A	B	C	D	E	F	G	H	I
1	Post	Reply	Score: 1 if YES. 0: if NO	Score: 1 if YES. 0: if NO	Score: 1 if YES. 0: if NO	Reply	Score: 1 if YES. 0: if NO	Score: 1 if YES. 0: if NO	Score: 1 if YES. 0: if NO
2			Grammatical	Coherent	Star Trek		Grammatical	Coherent	Star Trek
3	Beam me up Scotty.	Aye sir.	1	1	1	I do not have a teleportation device.	1	1	1
4	Dismissed.	I was just leaving this room.	1	1	0	M-5 is leaving itself open to attack. The machine is committing suicide to atone for the sin of murder.	1	0	1
5	Do you like me.	Bones I like you.	1	0	1	Sure, Lieutenant I like you very much.	1	1	1
6	Engage.	Warp one sir .	1	1	1	I am a Vulcan.	1	0	1
7	Hello	Hello Jim.	1	1	1	Yes, Lieutenant.	1	1	1
8	How are you.	Feeling fine.	1	1	0	Very well indeed, thank you, Lieutenant.	1	1	1
9	I am hungry captain.	What will you be eating?	1	1	0	I 'm glad you came to your ship .	1	0	1
10	I like you.	Bones i like you too	1	1	1	I like to make new friends.	1	1	0
11	I lost my phone.	Calm down Lieutenant.	1	1	1	Lieutenant.	0	0	1
12	Maximum power.	Remember, Lieutenant, what we see on this tape happened approximately three months ago.	1	0	1	Aye sir.	1	1	1
13	Mister Spock	Captain.	1	1	1	I am an artificial intelligence chatting on the Internet.	1	0	1
14	my name is alex	what do you want me to do Doctor	1	0	1	Yes, I know, Christine. Would you make me some of that plomeek soup?	1	0	1
15	red alert	Acknowledged. Taking emergency action.	1	1	1	Captain Uhura what ' s happening .	1	0	1
16	Shall i leave.	No Spock	1	0	1	Not right now, maybe later.	1	1	0
17	Spock to Captain.	This is ALICE	0	0	0	Captain Kirk	1	1	1
18	That is highly illogical.	I 'm glad you asked Lieutenant Kirk	1	0	1	Lieutenant, it also happens to be tape H, the impulse tracking obtained from the alien life units.	1	1	1
19	The message is being transmitted.	Picking up tricorder readings, Lieutenant. Apparently the ryetalyn is behind this door.	1	0	1	Captain .	1	0	1

# Demo





# Conclusion

- Our model able to automatically generate Star Trek styled text.
  - Even for out of domain input.
- Data driven approach. Can be easily expanded to other domains like news or sports.
- It can also be extended to emulate a specific fictional character.
- Future work:
  - Use more complex neural architectures in place of SEQ2SEQ.

**DO YOU HAVE**

**ANY QUESTIONS**

memegenerator.net