**Personality based chat-bot**

by

Ganesh Raj Kyatham

A Research Project Submitted in Partial Fulfillment of the Requirements

For the Degree of Master of Science in Statistics

Specialization in Data Science

At

Rutgers, The State University of New Jersey

March, 2024

### ABSTRACT

Chat-bots have become ubiquitous in various industries, serving diverse purposes from customer service to medical assistance. Despite their popularity, they still face several challenges, including maintaining context continuity, accurately recognizing user intents, and handling multi-turn dialogues. Additionally, Chat-bots struggle with emotion recognition, generating vague responses, and providing inconsistent answers. This project aims to address these issues by endowing Chat-bots with well-defined personalities. By ensuring greater consistency, tailoring responses to specific personalities, and introducing contextual awareness, the project seeks to enhance user experience and make Chat-bot interactions more seamless and human-like. Ultimately, the goal is to contribute to a more user-friendly and reliable Chat-bot experience.

## Table of Contents

1. Approach and Metrics 5
2. Dataset 6
3. Initial Experiments 7
4. Fine Tuned GPT2LMHead and XLNet 8
5. Seq2Seq with LSTM 12
6. Fine Tuned T5 14
7. Observations and Conclusion 15
8. Other Related work 16
9. References 18

**List of Tables**

1. Model Score Comparison 15

**List of Illustrations**

1. Screenshot of DataFrame 6

### APPROACH

This use case requires a text-to-text generation model that captures context information in a user prompt while also retaining information regarding the persona assigned to it through training.

Conversational models are the primary approach, although Q&A models are good at generating responses that are very specific and tailored to user input. For these reasons, this problem has been approached as both a conversational model problem and a Q&A model problem.

The modeling for this problem was done through the following stages:

1. **Initial Experiments with GPT-2 and Blenderbot**: This step informed us of the capability of pre-trained large language models to understand context and respond in the manner of a person just from the user prompt.
2. **Fine-Tuning GPT-2 and XLNet**: In this step, pre-trained large language models were finetuned using our dataset to make the responses more in line with the personality they were trained on. The models chosen for this step were GPT-2 LM Head and XLNet.
3. **Seq2Seq Model with LSTM**: In this step, a seq2seq model was created and trained on the problem dataset.
4. **Fine-Tuning T5-Small**: Since the seq2seq model wasn’t producing optimal results, reverting to pre-trained models seemed like a good approach. The final model chosen for this step was T5-Small.

**Evaluation Metrics**: BLEU score and ROUGE-1 F1, ROUGE-2 F1 scores were chosen as the key metrics to quantify and evaluate the model performances.

**BLEU Score**: BLEU Score, or Bilingual Evaluation Understudy, emerged as a crucial metric. It gauges the quality of generated text by evaluating the overlap of N-grams between the generated and reference texts.

**ROUGE Score**: The ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metric, comprising ROUGE-1 and ROUGE-2, was employed to evaluate unigram and bigram overlap, respectively. ROUGE-F1 combines both precision and recall providing a balanced metric.

Chat-bots typically respond to prompts with predefined answers, making BLEU and ROUGE ideal for evaluating their performance since these metrics compare generated text to reference responses. Both metrics gauge the relevance and context of the Chat-bot's replies, crucial for coherent and contextually appropriate conversations. BLEU's flexible N-grams and ROUGE-1's focus on unigrams offer a tailored assessment for Chat-bot dialogues.

The approach involved exploring both transformer and non-transformer models, fine-tuning techniques, and carefully selecting evaluation metrics to assess the Chat-bot's conversational and Q&A proficiency in natural language processing.

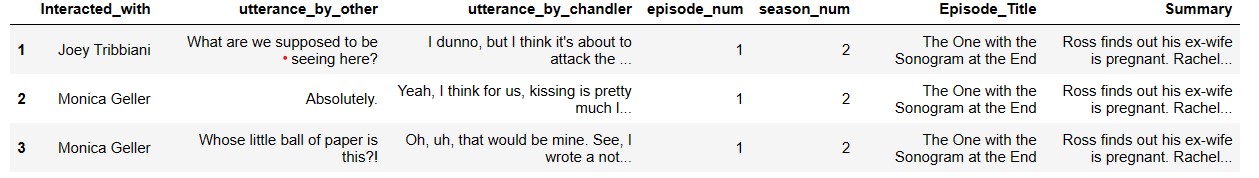
### DATASET

For this purpose, the FRIENDS CORPUS dataset, curated by Cornell University, was used owing to its easy access and comprehensive structure. The dataset included speaker-level, episodelevel, and scene-level data, along with some metadata regarding the scenes.

In total, the dataset encompasses approximately 80,000 dialogue interactions. For this specific use case, the interactions involving Chandler Bing were extracted, resulting in a total of 8568 pairs of dialogues.

Due to limitations in size and the absence of context for individual interactions within the dataset, the data was enhanced by adding the episode title and plot summary for each interaction, sourced from the FRIENDS plot summary dataset available on Kaggle. This new dataset includes titles and plot summaries for each episode, thereby introducing meaningful context.

Illustration -1 . Dataframe screenshot



To add context the title, plot summary and dialogue have been combined using small basic phrases and the final dialogue with context looks like this. The phrases used to join are highlighted in orange.

SAMPLE TEXT: in the episode The One with the Sonogram at the End the plot Ross finds out his ex-wife is pregnant. Rachel returns her engagement ring to Barry. Monica becomes stressed when her and Ross's parents come to visit. Chandler Bing said Oh, I think this is the episode of Three's Company where there's some kind of misunderstanding.

The dataset split ratio is 0.7 for training, 0.2 for validation, and 0.1 for testing.

### INITIAL EXPERIMENTS WITH GPT-2 and BLENDERBOT

Initial tests were conducted using the pre-trained large language models (LLMs) GPT-2 and Blenderbot. The objective was to assess their ability to comprehend detailed user prompts and respond in a specific manner, emulating the personality indicated in the user input. Blenderbot is a bidirectional transformer model that has 400 million parameters, while GPT2 is unidirectional, although is much bigger with 1.5 billion parameters.

The tokenizers and generate functions of these models were accessed through huggingface endpoints, that enabled to use these functions without having to loading the entire models on local device.

The outcomes for sample prompts were as follows:

**Text Prompt:** In the episode "The One with Two Parts: Part 1," the plot involves Phoebe feeling neglected when Joey falls for her twin sister. Chandler struggles to fire an attractive co-worker. Ross attends Lamaze classes with Carol and Susan. Ursula said Hi. Okay, will that be all?

* **True output:** Chandler said: Wait, wait! Wh-what are you doing here?
* **BlenderBot Output:** I'm not sure, but I do know that the plot revolves around two brothers.
* **GPT2 Output:** In the episode "The One with Two Parts: Part 1," Phoebe feels neglected when Joey falls for her twin sister. Chandler struggles to fire an attractive co-worker. Ross attends Lamaze classes with Carol and Susan. Ursula said Hi. Okay, will that be all? Ph...

**Text Prompt:** In the episode "The One with All the Poker," the plot revolves around Rachel interviewing for a new job. The girls take on the guys in a game of poker. Ross Geller said, "I don't know." Chandler said, "Is this still about her whole 'The Flintstones could've really happened' thing?"

* **True output:** Chandler said: Is this still about her whole 'The Flintstones could've really happened' thing?
* **BlenderBot Output:** Chandler is a character in the animated series "The Flintstones," which was created by Hanna-Barbera.
* **GPT2 Output:** In the episode "The One with All the Poker," the plot As Rachel interviews for a new job, the girls take on the guys in a game of poker. Ross Geller said, "I don't know." Chandler said, "Is this still about her whole 'The Flintstones could've really happened...

**Text Prompt:** In the episode "The One with the Fake Monica," the plot involves Monica befriending the woman who stole her credit card. Ross is forced to give away Marcel. Joey tries to come up with a less ethnic-sounding stage name. Ross Geller said, "I think when someone steals your credit card, they've kind of already thrown caution to the wind."

* **True output:** Chandler said: Wow, what a geek. They spent $69.95 on a Wonder Mop.
* **BlenderBot Output:** That's a good way to look at it. I can't believe that the show has been on for so long. It premiered in 1997 on NBC.
* **GPT2 Output:** In the episode "The One with the Fake Monica," the plot Monica befriends the woman who stole her credit card. Ross is forced to give away Marcel. Joey tries to come up with a less ethnic-sounding stage name. Ross Geller said, "I think when someone steals your credit card, they've kind of already thrown caution to the wind”. Ross.

**Conclusion:** As these models are fundamentally general-purpose text-to-text generators, they struggled to consistently mirror the desired personality from a single prompt. Additionally, the lack of detailed knowledge about the subject's personality in the models led to responses that didn't fully meet our expectations.

### FINE TUNED GPT-2LMHEAD

GPT-2, or the "Generative Pre-trained Transformer 2 is a generative transformer model that uses self-attention to capture context and long-range dependencies in the data. GPT2-LMHead is a version of GPT2 that has an additional linear head for language modelling. This model has been trained on outgoing webpages from Reddit and has 124M parameters. This model has been accessed through the `transformers` python library by Huggingface.

This was implemented as an unsupervised model.

**Input Data Format:**

<sos>inputdata <bot> outputdata<eos> . Here <sos>,<bot>, <eos> are special tokens that indicate Start of Sentence, Bot text generation prompt token and end of sentence respectively.

**Model Parameters:**

Since this is a pretrained model, there were only a handful of parameters to tune.

max\_new\_tokens = 50, weight\_decay = 0.2, epochs = 15 Weight decay was increased to 0.2 to avoid overfitting.

**Sample Output:**

1. **Input Prompt: In the episode** "The One with Two Parts: Part 1," the plot involves Phoebe feeling neglected when Joey falls for her twin sister. Chandler struggles to fire an attractive coworker. Ross attends Lamaze classes with Carol and Susan. Ursula said Hi. Okay, will that be all?

**Actual Output: Chandler said:** Wait, wait! Wh-what are you doing here?

**Model Output:** ['in', 'the', 'episode', 'the', 'one', 'with', 'two', 'parts', 'part', '1', 'the', 'plot',

'phoebe', 'feels', 'neglected', 'when', 'joey', 'falls', 'for', 'her', 'twin', 'sister', '.', 'chandler', 'struggles', 'to', 'fire', 'an', 'attractive', 'co-worker', '.', 'ross', 'attends', 'lamaze', 'classes', 'with',

'carol', 'and', 'susan', '.', 'ursula', 'said', 'hi', '.', 'okay', ',', 'will', 'that', 'be', 'all', '?', 'chandler', 'said', 'wait', ',', 'wait', '!', 'wh', '-', 'what', 'are', 'you', 'doing', 'here', '?']

1. **Input Prompt: In the episode** "The One with All the Poker," the plot revolves around Rachel interviewing for a new job. The girls take on the guys in a game of poker. Ross Geller said, "I don't know." Chandler said,"Is this still about her whole 'The Flintstones could've really happened' thing?"

**Actual Output:** Chandler said:Is this still about her whole 'The Flintstones could've really happened' thing?

**Model Output:** ['in', 'the', 'episode', 'the', 'one', 'with', 'all', 'the', 'poker', 'the', 'plot', 'as',

'rachel', 'interviews', 'for', 'a', 'new', 'job', ',', 'the', 'girls', 'take', 'on', 'the', 'guys', 'in', 'a', 'game',

'of', 'poker', '.', 'ross', 'geller', 'said', 'i', 'don', "'", 't', 'know', '.', 'chandler', 'said', 'is', 'this', 'still',

'about', 'her', 'whole', "'", 'the', 'flintstones', 'could', "'", 've', 'really', 'happened', "'", 'thing', '?']

1. **Input Prompt: In the episode** "The One with the Fake Monica," the plot involves Monica befriending the woman who stole her credit card. Ross is forced to give away Marcel. Joey tries to come up with a less ethnic-sounding stage name. Ross Geller said, "I think when someone steals your credit card, they've kind of already thrown caution to the wind."

**Actual Output:** Chandler said: Wow, what a geek. They spent $69.95 on a Wonder Mop.

**Model Output:** ['in', 'the', 'episode', 'the', 'one', 'with', 'the', 'fake', 'monica', 'the', 'plot',

'monica', 'befriends', 'the', 'woman', 'who', 'stole', 'her', 'credit', 'card', '.', 'ross', 'is', 'forced',

'to', 'give', 'away', 'marcel', '.', 'joey', 'tries', 'to', 'come', 'up', 'with', 'a', 'less', 'ethnic-sounding',

'stage', 'name', '.', 'ross', 'geller', 'said', 'i', 'think', 'when', 'someone', 'steals', 'your', 'credit',

'card', ',', 'they', "'", 've', 'kind', 'of', 'already', 'thrown', 'caution', 'to', 'the', 'wind', '.', 'chandler', 'said', 'wow', ',', 'what', 'a', 'geek', '.', 'they', 'spent', '$', '69.95', 'on', 'a', 'wonder', 'mop']

**Conclusion:** Although GPT2 is a model that has been trained on large datasets, it fails to adapt to this specific task using the small dataset that was provided. The model simply repeated the target text and failed to generate any meaningful or relevant output. It did generate one or two barely decent inputs during the training, but we could not replicate those results in the sample prompts we tried.

### FINE TUNED XLNet

Like GPT-2, XLNet is an autoregressive transformer, but it has an additional feature: it is bidirectional. It captures bidirectional context and dependencies effectively, allowing it to understand intricate relationships in the input text. The base-cased model has 110M parameters, while the large-cased model is 340M parameters strong.

**Input Data Format:**

The format of the input is the same as for GPT-2, with the prompt and response together forming a sentence, but separated and bound by the afore-mentioned special tokens.

**Model Parameters:**

Following are the final values of the model parameters that we could fine-tune.

* max\_new\_tokens = 50
* weight\_decay = 0.1
* epochs = 1

**Sample Output:**

*In-sample:*

**Text prompt:** What? You You Oh! Can I ask you a personal question? Ho-how do you shave your beard so close?

* **True response:** Ok Tommy, that's enough mourning for you! Here we go, bye bye!!
* **Generated response:** <startofstring>What? You You Oh! Can I ask you a personal question? Ho-how do you shave your beard so close?<bot>:<sep><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls ><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls>< cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls ><cls><cls>

**Text prompt:** Uh, we might be leaving now.

* **True response:** Tell me it's "you and me" we.
* **Generated response:** <startofstring>Uh, we might be leaving now.<bot>:<sep><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls> <cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cl s><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls> <cls><cls>

**Text prompt:** I know what time you said.

* **True response:** What?!
* **Generated response:** <startofstring>I know what time you said.<bot>:<sep><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls>< cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls ><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls>< cls><cls>

*Out-of-sample:*

**Text prompt:** In the episode The One with Two Parts: Part 1 the plot Phoebe feels neglected when Joey falls for her twin sister. Chandler struggles to fire an attractive co-worker. Ross attends Lamaze classes with Carol and Susan. Ursula said Hi. Okay, will that be all?

* **True response:** chandler said Wait, wait! Wh-what are you doing here?
* **Generated response:** <startofstring>in the episode The One with Two Parts: Part 1 the plot Phoebe feels neglected when Joey falls for her twin sister. Chandler struggles to fire an attractive co-worker. Ross attends Lamaze classes with Carol and Susan. Ursula said Hi. Okay, will that be all?<bot>:<sep><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><c ls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls> <cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cl s><cls>

**Text prompt:** In the episode The One with All the Poker the plot As Rachel interviews for a new job, the girls take on the guys in a game of poker. Ross Geller said I don't know.chandler said chandler said chandler said Is this still about her whole 'The Flintstones could've really happened' thing?

* **True response:** chandler said Is this still about her whole 'The Flintstones could've really happened' thing?
* **Generated response:** <startofstring>in the episode The One with All the Poker the plot As Rachel interviews for a new job, the girls take on the guys in a game of poker. Ross Geller said I don't know.chandler said chandler said chandler said Is this still about her whole 'The Flintstones could've really happened' thing?<bot>:<sep><cls><cls><cls><cls><cls><cls>chandler said chandler said chandler said chandler said chandler said chandler said chandler said chandler said chandler said chandler said chandler said

**Text prompt:** in the episode The One with the Fake Monica the plot Monica befriends the woman who stole her credit card. Ross is forced to give away Marcel. Joey tries to come up with a less ethnic-sounding stage name. Ross Geller said I think when someone steals your credit card, they've kind of already thrown caution to the wind.

* **True response:** chandler said chandler said Wow, what a geek. They spent $69.95 on a Wonder Mop.
* **Generated response:** <startofstring>in the episode The One with the Fake Monica the plot Monica befriends the woman who stole her credit card. Ross is forced to give away Marcel. Joey tries to come up with a less ethnic-sounding stage name. Ross Geller said I think when someone steals your credit card, they've kind of already thrown caution to the wind.<bot>:<sep><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls> <cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cl s><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls><cls> <cls><cls>

**Conclusion:** Clearly, the model is overfitted on the input and the special tokens. The small size of the dataset exacerbates this problem. The same problem was observed during the training phase as well.

### SEQ2SEQ model with LSTM

Since Seq2Seq models are built using RNNs as base units, they adeptly handle input of varying lengths. The encoder captures the essence of the input, and through the final layer, it passes this context vector as input to the decoder. To mitigate the problem of vanishing gradients, the Seq2Seq model uses LSTM as the base unit. LSTMs utilize a memory cell and a set of gates to selectively store and forget data, enabling them to capture long-term dependencies. This is a supervised model.

The details of the Seq2Seq model that have been used are mentioned below:

**Input Data Format:**

The source sentence is the dialogue with context.

The target sentence is the Dialogue uttered by Chandler.

**Model parameters:**

Encoder embedding size = 250, Decoder embedding size = 250, Number of layers = 2, Encoder dropout = 0.5, Decoder dropout = 0.5

The loss metric used was cross entropy loss.

To prevent overfitting, dropout has been added to both encoder and decoder units.

**Sample output:**

**Input Prompt:** In the episode "The One with Two Parts: Part 1," the plot involves Phoebe feeling neglected when Joey falls for her twin sister. Chandler struggles to fire an attractive co-worker.

Ross attends Lamaze classes with Carol and Susan. Ursula said Hi. Okay, will that be all?

* **True output:**
* Chandler said: Wait, wait! Wh-what are you doing here?
* **Model Output:**
* ['in', 'the', 'episode', 'the', 'one', 'with', 'all', 'the', 'resolutions', 'the', 'plot', 'the', 'gang',

'constantly', 'help', 'for', 'her', 'visit', 'staff', 'to', 'acknowledge', '.', 'phoebe',

'temporarily', 'moves', 'into', 'their', 'apartment', 'and', 'phoebe', 'are', 'attracted', 'to',

'her', 'dates', 'a', 'stage', 'resemblance', 'to', 'kathy', '.', 'rachel', 'said', 'said', 'yeah', '.', 'well', ',', 'what', "'s"]

**Input Prompt:** In the episode "The One with All the Poker," the plot revolves around Rachel interviewing for a new job. The girls take on the guys in a game of poker. Ross Geller said, "I don't know." Chandler said, "Is this still about her whole 'The Flintstones could've really happened' thing?"

* **True output:**
* Chandler said: Is this still about her whole 'The Flintstones could've really happened' thing?
* **Model Output:**
* ['in', 'the', 'episode', 'the', 'one', 'with', 'all', 'the', 'poker', 'the', 'plot', 'as', 'rachel',

'interviews', 'for', 'a', 'new', 'job', ',', 'the', 'girls', 'take', 'on', 'the', 'guys', 'in', 'a', 'game',

'of', 'poker', '.', 'chandler', 'bing', 'said', 'bladder', 'control', 'problem', '?', '<unk>',

'<unk>', '<unk>', '<unk>', '<unk>', '<unk>', '<unk>', '<unk>', '<unk>', 'a', '<unk>', '<unk>']

**Input Prompt:** In the episode "The One with the Fake Monica," the plot involves Monica befriending the woman who stole her credit card. Ross is forced to give away Marcel. Joey tries to come up with a less ethnic-sounding stage name. Ross Geller said, "I think when someone steals your credit card, they've kind of already thrown caution to the wind."

* **True output:**
* Chandler said: Wow, what a geek. They spent $69.95 on a Wonder Mop.
* **Model Output:**
* ['in', 'the', 'episode', 'the', 'one', 'with', 'the', 'fake', 'monica', 'the', 'plot', 'monica',

'befriends', 'the', 'woman', 'who', 'stole', 'her', 'credit', 'card', '.', 'ross', 'is', 'forced', 'to', 'give', 'away', 'marcel', '.', 'joey', 'tries', 'to', 'come', 'up', 'with', 'a', 'less', 'ethnicsounding', 'stage', 'name', '.', 'phoebe', 'buffay', 'said', 'the', 'he', 'is', '<unk>', 'the', '<unk>']

**Conclusion:** For the first prompt, some parts of the output were coherent like “Phoebe temporarily moves into their apartment”, although it is still not relevant to Chandler Bing’s response. The model tried to generate the most popular tokens in the dataset at first, but after those tokens, it sometimes did generate meaningful output. This model performed decently considering this isn’t a pre-trained model and it has only been trained on a small dataset.

### T5 fine-tuned model

The Text-to-Text Transfer Transformer, or T5, is a versatile language model that tackles various natural language processing tasks in a unified manner by framing them as text-to-text problems. It is a transformer model that uses self-attention mechanisms to capture complex contexts in the data. T5-small, a much smaller version of T5 with 60 million parameters, has been used for fine-tuning in this use case. This is a supervised model.

The T5 model has been trained on several input prefixes that indicate the type of task it is supposed to perform. For this use case, the prefix used is 'chat:’, after testing several others such as 'respond to:', 'reply:', and 'converse:', 'chat' seemed to yield the best results.

The model was downloaded through transformers library from hugging-face-hub-repository and trained using Seq2seq trainer.

More details about this model are as follows:

**Input data format:**

The source sentence is the dialogue with the context with the prefix “chat:” The target sentence is the dialogue uttered by Chandler.

**Model parameters:**

*Training parameters:*

Since this is a fine-tuned model there were only a handful of a parameters available to tune.

Number of epochs = 5, Learning rate = 2e-5, Batch size = 16, Weight decay = 0.05. Weight decay of 0.05 was added to prevent the model from repeating the user prompt.

*Generation parameters:*

Max length = 50, Temperature = 0.9, Repetition penalty = 1.9, Number of beams = 5, No repeat n gram size = 2

Temperature was raised to 0.9 to introduce randomness to the responses. Max length was set to 50, so that the responses are too long where context can be lost. Number of beams is set to 5 to enable more combinations of output tokens.

**Sample output**:

**Input Prompt 1:** Chat: In the episode "The One with Two Parts: Part 1," the plot involves

Phoebe feeling neglected when Joey falls for her twin sister. Chandler struggles to fire an attractive co-worker. Ross attends Lamaze classes with Carol and Susan. Ursula said Hi. Okay, will that be all?

* **True output 1:** Chandler said: Wait, wait! Wh-what are you doing here?
* **Model Output 1:** Chandler said: Okay, will that be all?

I**nput Prompt 2:** Chat: In the episode "The One with All the Poker," the plot revolves around

Rachel interviewing for a new job while the girls take on the guys in a game of poker. Ross Geller said, "I don't know." Chandler said, "Is this still about her whole 'The Flintstones could've really happened' thing?"

* **True output 2:** Chandler said: Is this still about her whole 'The Flintstones could've really happened' thing?
* **Model Output 2:** Chandler said: Is this still about her whole 'The Flintstones could've really happened' thing?

**Input Prompt 3:** Chat: In the episode "The One with the Fake Monica," the plot involves Monica befriending the woman who stole her credit card. Ross is forced to give away Marcel. Joey tries to come up with a less ethnic-sounding stage name. Ross Geller said, "I think when someone steals your credit card, they've kind of already thrown caution to the wind."

* **True output 3:** Chandler said: Wow, what a geek. They spent $69.95 on a Wonder Mop.
* **Model Output 3:** Chandler said: I think when someone steals your credit card, they've kind of already thrown caution to the wind.

**Conclusion:** The model being trained on very large corpus, has generated coherent output in some cases but it repeated the true output in rest of the cases. However, it still didn’t generate output in the manner desired by this use case.

### Observations and Conclusion

Table – 1 Model score comparisons

|  |  |  |  |
| --- | --- | --- | --- |
| **MODEL** | **BLEU SCORE** | **ROGUE 1 F1** | **ROGUE 2 F1** |
| GPT2 | 0.63 | 0.89 | 0.85 |
| XLNet | 0.71 | 0.94 | 0.93 |
| SEQ2SEQ | 0.11 | 0.32 | 0.16 |
| T5 | 0.063 | 0.17 | 0.01 |

* While the T5 model generated sensible outputs, the metrics indicated poor performance, highlighting potential limitations in the evaluation metrics used.
* These metrics mainly focus on matching the generated output with the true output and might not fully capture the effectiveness of the generated responses in certain contexts.
* Crucial aspects of Chandler Bing's personality, such as humor and style, are not measured by these metrics, potentially leading to an underestimation of the model's performance in replicating nuanced character traits.
* The relatively small size of the dataset could also be a limiting factor; a larger dataset might be necessary for the model to accurately learn and replicate a complex personality using dialogues.
* Additionally, in the dataset, conversations were treated as individual data points, but reallife conversations have a continuous flow. This approach may affect the model's learning and output.

**Future scope:**

* Implement seq2seq model with attention mechanism that will better capture context and long-term dependencies in the dataset.
* Feed the model with continuous and complete conversation instead of passing individual dialogue pairs.
* Implement encoding the interactions for each person in the train dataset.

**Useful papers inferred:**

Here are a few papers have contributed innovative insights to this topic:

* A Persona-Based Neural Conversation Model. Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, Bill Dolan arXiv:1603.06155

This paper implemented SEq2seq lstm frameworks using two dialogue models: single-speaker model and dyadic speaker-addresse model.

Twitter Persona dataset was used for implementing the models. The model was trained on IMSDb(internet movie script database) and then the pretrained model has been fine tuned using the twitter persona dataset. The innovative thing about this research paper was the creation and use of individual embeddings for each personality in the single-speaker model caseand embeddings for all pairwise speaker-addressee combinations and passing them to the decoder along with the input from the final step of the encoder. This assigns each conversation to a specific personality or a combination of addressee-speaker pair. The final models were successful in capturing certain personal characteristics such as personal talking style and background information. The fine tuned models also had a better performance than the seq2seq models directly trained on twitter persona dataset.

* I am Kalam” - Analyzing and Generating Kalam's Answer Patterns

This paper created a seq2seq model that would respond to queries in the talking style of Dr. Abdul Kalam.For the dataset, Abdul Kalam's quotes from various books and journals that were available on Goodreads were extracted. The quotes are already tagged to the topics they are related to, which could be geography, science, politics ..etc. Abdul Kalam's responses to questions were extracted too from goodreads. A Q&A model was created using seq2seq architecture that was trained on the dataset.

From the prompt responses in the paper, the model seemed to have successfuly replicated Dr Kalam's answering style to some extent.

**References:**

* Alexander Castañeda, Patrick Brown, Rais Kazi, Landyn Moreno, Christian Tomah, Phillip Peng, Michael Hildner GPT-3 Trained To Impersonate By: Alexander Castañeda
* A Persona-Based Neural Conversation Model. Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, Bill Dolan arXiv:1603.06155
* Personalizing Dialogue Agents: I have a dog, do you have pets too? Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, Jason Weston
* TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents by Thomas Wolf, Victor Sanh, Julien Chaumond & Clement Delangue
* Towards Empathetic Open-domain Conversation Models: a New Benchmark and Dataset by Hannah Rashkin, Eric Michael Smith, Margaret Li, Y-Lan Boureau
* Chat-bots with Personality Using Deep Learning by Gaikwad, Susmit (2019). Master's Projects. 678
* Automatic Evaluation of Neural Personality-based Chat-bots by Xing, Y. and Fernández, R. (2018)
* I am Kalam – Analyzing and Generating Kalam’s Answer Patterns