
Code Slingers Project Idea

— Language Generation Tasks, SoSe2021 —

Isidora Jeknic, Iuliia Zaitova, Kirstin Kolmorgen, Sharmila Upadhyaya

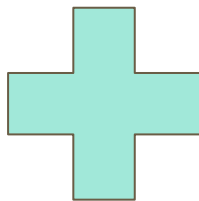
Outline

1. Introduction
2. Background
3. Novelty
4. Dataset
5. Architecture
6. Timeline
7. Expected Result

Introduction

Project idea: Combine image captioning with joke generation

Image Captioning



Joke Generation

Why don't seagulls live by the bay?
Because they would be bagels.

Background

Template-based systems: model selects words and fills them into fixed slots

- **Light Bulb Joke Generator** (LIBJOG, Raskin and Attardo, 1994):

‘How many <group name> does it take to screw in a light bulb?

<NumberX>. One to <activity1> and <NumberY> to <activity2>.’ →

‘How many Californians does it take to screw in a light bulb? Twelve. One to screw it in and eleven to share in the experience.’

- **Petrovic and Matthews** (2013):

‘I like my X like I like my Y, Z.’ →

‘I like my relationship like I like my source, open.’

Background

Neural Systems: more creative, but mostly fail to generate funny jokes

- **Yang and Ren** (2017): LSTM trained on joke corpus; user can provide topics

Topic Words	Generated Joke (Selected)
Apple Playboy	Apple is teaming up with Playboy Magazine in the self driving office .
None	One of the top economy in China , Lady Gaga says today that Obama is legal .

Background

- Yu et al. (2018): encoder-decoder framework trained on Wikipedia corpus;
input: a polysemic word and two of its meanings

pitch: 1) the property of sound that arise with variation in the frequency of vibration;
2) the act of throwing a baseball by a pitcher to a batter.

Highlight	in one that denotes player may have had a high pitch in the world
Joint	the object of the game is based on the pitch of the player

Novelty

The *synthesis* of image captioning and humor generation

- joke generation based on entity extraction
- combination of formulaic generation (c.f. template-based), and user-input generation (c.f. neural-based)
- novel architecture (more on that later)
- web app
- actually funny output

Dataset

A subset of the “*Short Jokes*” dataset (231,657 jokes ranging from 10 to 200 characters):

- consists of jokes that include a **why** question (16553, pre-filtering); format:

[Why] + [NP] + [VP]?
- (Because) [JOKE].

-
- offensive jokes need to be filtered out

Link to dataset: <https://www.kaggle.com/abhinavmoudgil95/short-jokes>

Dataset

- examples (*why*):

Why don't mathematicians build bridges? Because they'll never be perfect.

Why did the greek philosopher break in two? He was made out of Plato

Why did the boy die? Because he ate honey nectar and then went to the store and tried to buy some mustard and got eaten by a Stegosaurus

Why did black beard kill his first mate? He kept singing, "your ho, your ho, a pirate's wife for me"

Why couldn't the beaver cook his twigs? He only had Non-Stick pans.

Why are Fencers so popular on /r/Jokes? Because they do well with Ripostes.

Why Didn't The Skeleton Go To The Party? Because he had no body to go with

Why did the Fonz have red eyes? Because he had ****AYYY**** fever!

Why does Santa's sled get such good mileage? Because it has long-distance runners on each side.

Why did Plato's worst student always have stuff stuck in his teeth? He didn't understand Flossophy!

Why did the chicken run out into traffic? To get to the other side.

Why did the cow sign up for TSA precheck? Butterflies

Dataset

- could be expanded to include:

1. **what** jokes (pre-filtering: 41659)

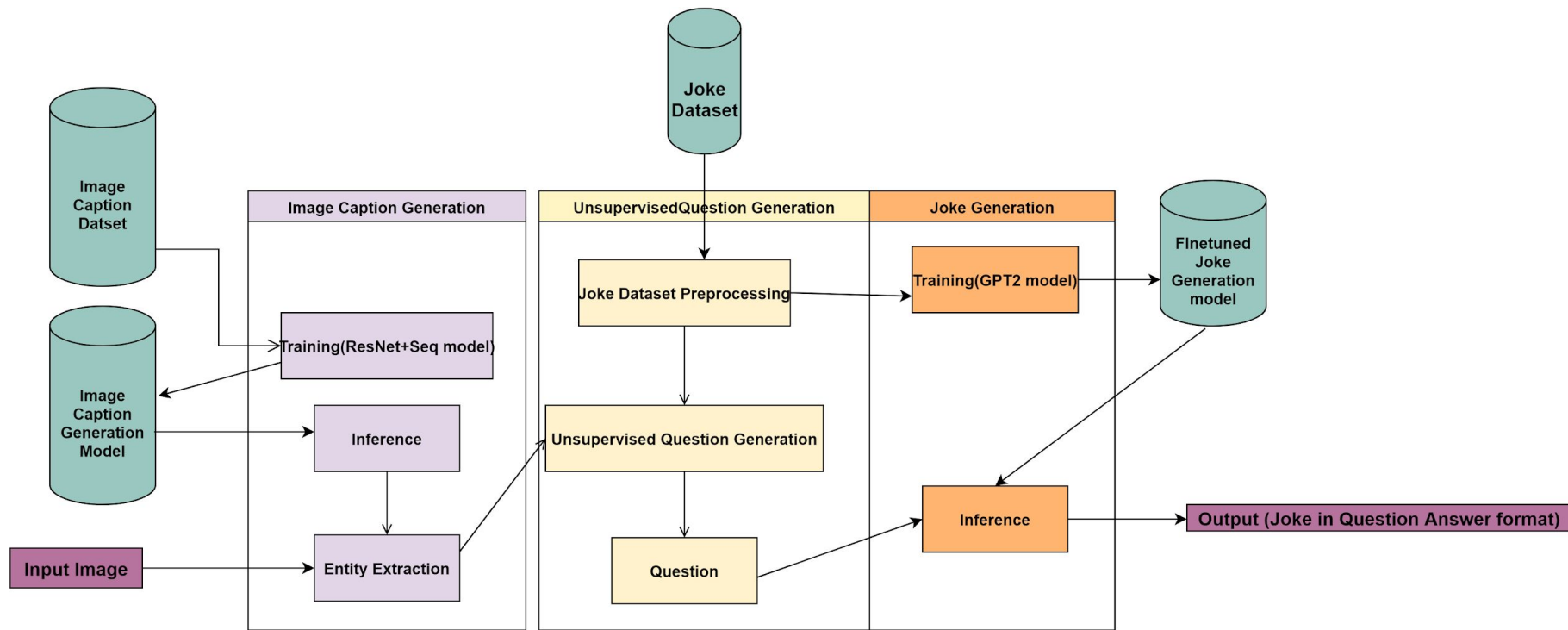
What did the pre-school math teacher have to say about her students? Every single one counts.
What do you call a crushed angle? A rektangle.
What kind of dog is the smartest? A great brain!
What is the gender of Iron Man? Fe Male
What kind of food can you color with? A cranberry! ~ Esher (my Grandson) age 5
What did the conceited man say while he stood on the north pole? The earth revolves around me.
What's a hipster's favorite kind of cigarette? Yours.
What did the duck say to the cashier after she purchased some lipstick? "Just put it on my bill"
What would Marilyn Monroe be doing if she were alive today? Clawing at the inside of her coffin.
What's a pirates favorite letter? You may think R, but a pirates first love will always be the C.
Whats a caged animal's favorite food? Zoocchini
What do you call a smart mexican? Sherlock, Holmes.
What is M. Night Shamalan's favorite game? Twister.
What is Hitler's favorite drink? genocider

Dataset

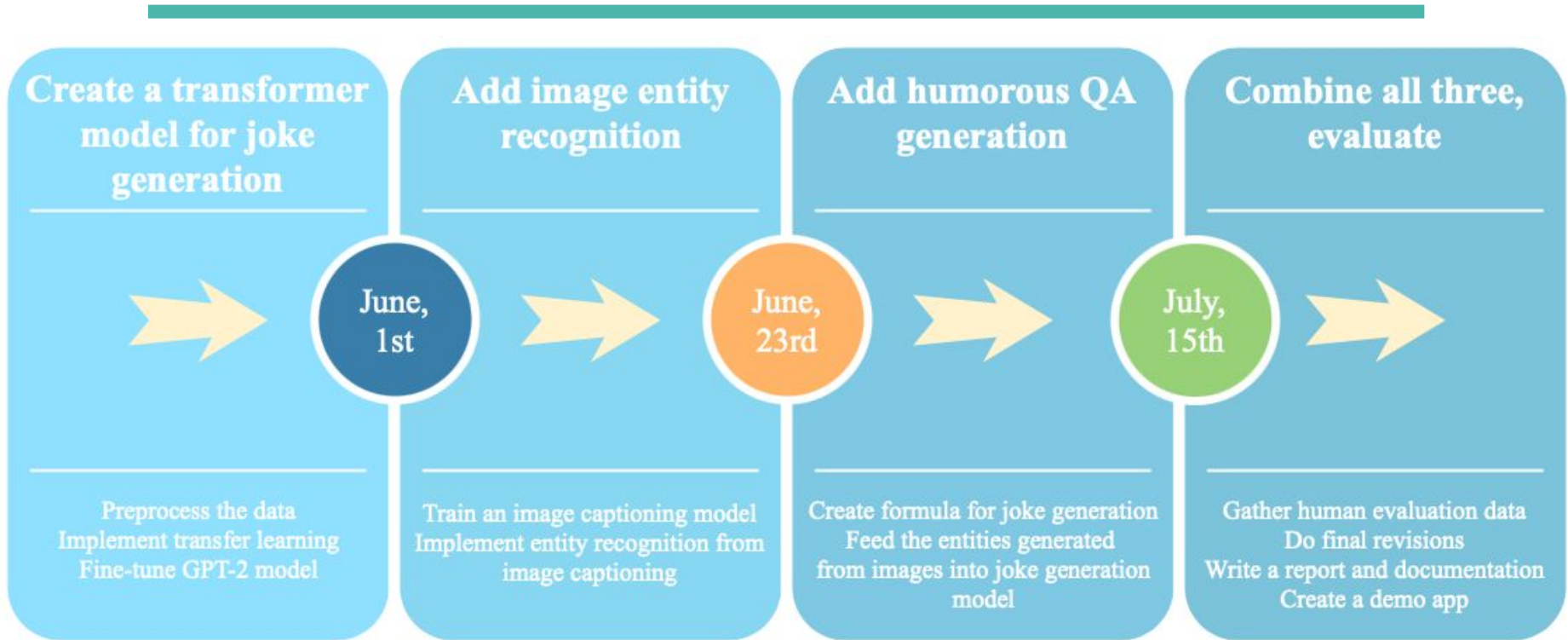
2 *how* jokes (pre-filtering: 10642)

How do you call a blond who changed her hair color to brown? Artificial Intelligence!
How does software eat it's food? By taking large bytes!
How does every racist joke begin? By glancing around the room.
How do you unlock a door made of dough With a cookey
How much for that babysitter? Ma'am, that's a roll of duct tape I'll take it!
how do you make pickle bread? with dill dough
How do you circumsize an Alabama man? Kick his sister in the chin.
How many bees does it take to screw in a light bulb? Two, but how do they get in there?
How do skeletons make a baby? They bone
How do you know if a guy has an asian wife? He'll tell you.
How many bassists does it take to screw in a lightbulb? None the keyboardist can do it with his left hand.
How long does it take for a GNOME developer to take a shit? I don't know, how long until the next version of GNOME?
How many men does it take to close the toilet seat? No one knows yet!
How would you tell if the Queen was stoned? Her Highness will tell you

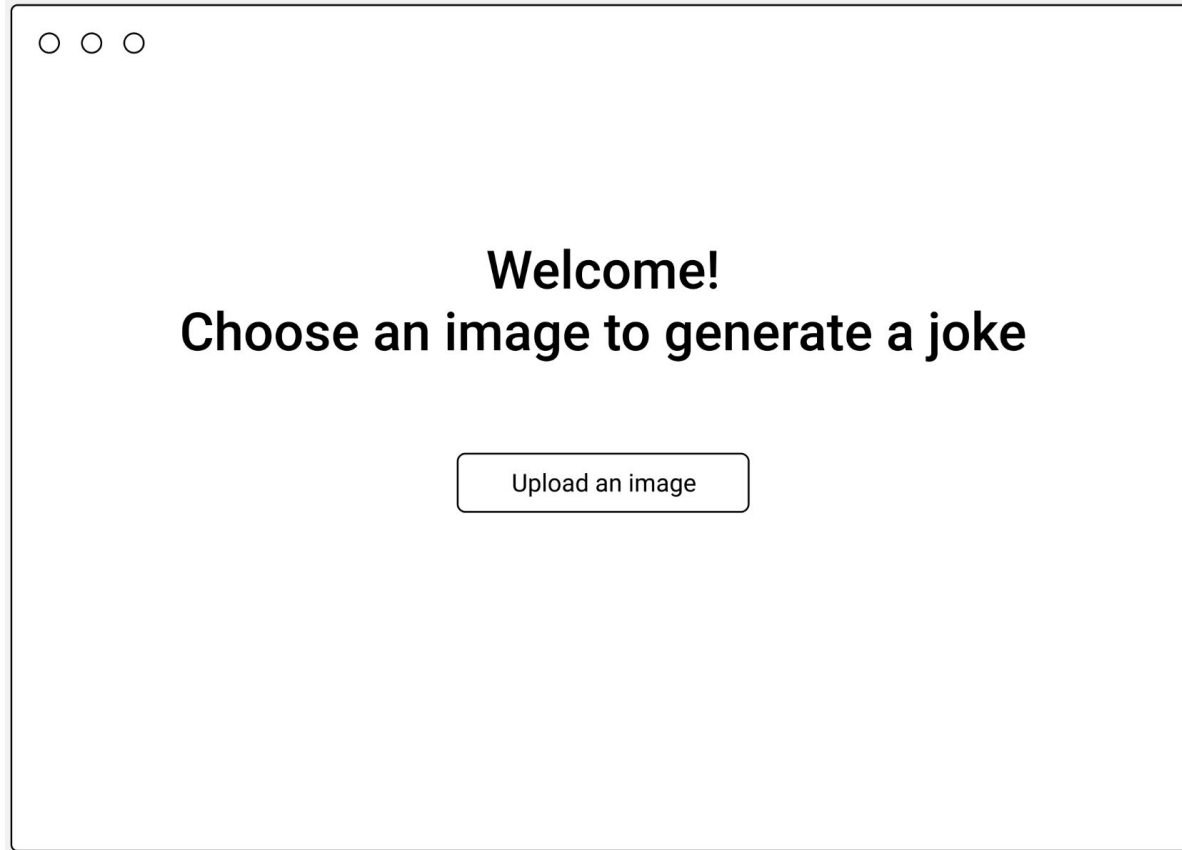
Architecture



Timeline



Expected Results



Expected Results



Why do rabbits love beer? Because it's made of hops.

Expected Results

Lack of accuracy

<i>Source</i>	<i>4+ ratings</i>
GAG (Classifier)	11.41%
GAG (Regression)	10.12%
Human (All)	27.38%
Human (JokeJudger)	22.61%
Human (Single words)	21.08%

Automatic Joke Generation: Learning Humor from Examples

Thomas Winters, Vincent Nys, and Daniel De Schreye

Model	Human Evaluation Score	% of Good Jokes (Score = 2, 3)
Human Jokes	2.12	35.3%
Probabilistic Model in [3] (baseline)	1.44	16.3%
Neural Joke Generator	1.59	12.1%
<i>Score: 0 - nonsense; 1 - not funny; 2 - somewhat funny; 3 - funny</i>		

Neural Joke Generation

He Ren, Quan Yang

Questions?

Thank you for your attention!

References

- Petrović, S., & Matthews, D. (2013). *Unsupervised joke generation from big data* (pp. 228–232). Association for Computational Linguistics.
<https://www.aclweb.org/anthology/P13-2041.pdf>
- Raskin, J. D., & Attardo, S. (1994). Non-literality and non-bona-fide in language: An approach to formal and computational treatments of humor. *Pragmatics and Cognition*, 2(1), 31–69. <https://doi.org/10.1075/pc.2.1.02ras>
- Ren, H., & Yang, Q. (2017). *Neural Joke Generation*. <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1174/reports/2760332.pdf>
- Winters, T., Nys, V., & De Schreye, D. (2018). Automatic Joke Generation: Learning Humor from Examples. *Distributed, Ambient and Pervasive Interactions: Technologies and Contexts*, 360–377. https://doi.org/10.1007/978-3-319-91131-1_28
- Yu, Z., Tan, J., & Wan, X. (2018). *A Neural Approach to Pun Generation* (pp. 1650–1660). Association for Computational Linguistics.
<https://www.aclweb.org/anthology/P18-1153.pdf>