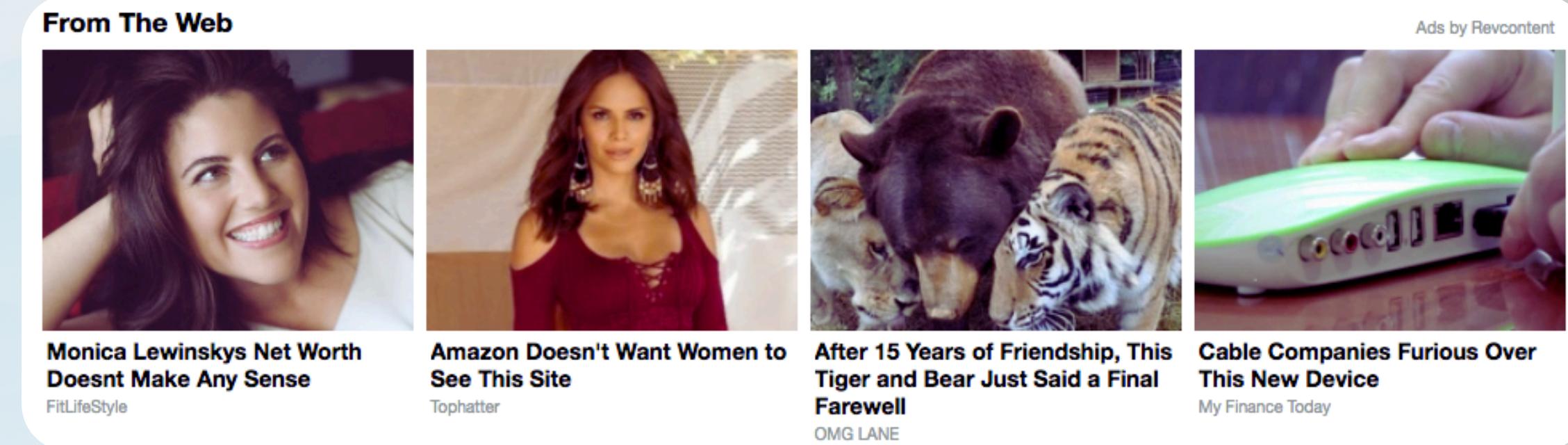


Clickbait Spoiler Detection and Generation

Poluri Rishi Manoj (CS21BTECH11045)

Kota Vignan (CS21BTECH11029)

- Group 17



Problem Statement

- Clickbaits are texts containing sensationalized terms that aim at raising the curiosity of the readers, and lure them to click on a hyperlink. These are common in news headlines, social media posts etc.
- These usually fall short of readers' expectations, wasting their time and causing disappointment.

Due to the deceptive and misleading nature of the content often presented, clickbait is generally considered harmful. Thus addressing the clickbait issue is a crucial concern.

Solution:

- The generation of spoilers, i.e. short texts that neutralize clickbait by providing information that satisfies the curiosity induced by it.

Spoilers

- Spoilers refer to short summaries or key pieces of information that reveal the core content behind a sensationalized or misleading headline, eliminating the need for users to click through to the full article.
- The goal is to "spoil" the bait of the clickbait by giving readers the factual information upfront.

Types of Spoilers:

- phrase: Spoiler of single word or phrase
 - Ex: a nap
- passage: Spoiler composed from continuous sentences
 - Ex: Highly skilled and in-demand work at a high salary will increase your assets.
- multi: Spoiler comprised of multiple discontinuous words or sentences
 - 1. Exercise
 - 2. Eat balanced diets
 - 3. getting enough sleep

Webis Clickbait Spoiling Corpus 2022

– Dataset Overview

- Webis-Clickbait-22 contains 5,000 spoiled clickbait posts crawled from Facebook, Reddit, and Twitter.
- This dataset contains the clickbait posts and manually cleaned versions of the linked documents, and extracted spoilers for each clickbait post.
- Data Fields: uuid, postId, postText, postPlatform, targetParagraphs, targetTitle, targetDescription, targetKeywords, targetMedia, targetUrl, provenance, spoiler, spoilerPositions, tags.
- Datasplit:
 - train: 3,200 posts
 - validation: 800 posts
 - test: 1000 posts
- Proportion of Spoiler Types in training data:
 - 1367 (around 43%) for phrase
 - 1274 (around 40%) for passage
 - 559 (around 17%) for multi

Examples

- {"uuid": "d636e94d-d27a-4366-b177-1299c266d7b5"}
- "postId": "553676574290354177"
- "postText": ["The anytime snack you won't feel guilty about eating"]
- "postPlatform": "Twitter"
- "targetParagraphs": ["Whether you're between meals or craving a midnight snack, there's no need to feel guilty about getting your fill. This healthy, spiced snack is lower in calories and fat than other go-to options, like nuts, and is sure to leave you feeling more than satisfied.", "Smoky Paprika-Baked Garbanzo Beans", "Recipe created by Art Smith", "Serves 4", "Ingredients", "1 tsp. onion powder", "1 Tbsp. smoked paprika", "30 oz. canned garbanzo beans, drained", "Directions", "Preheat the oven to 350 degrees.", "Pour the drained garbanzo beans into a medium mixing bowl with the onion powder and smoked paprika. Mix well.", "Spread the coated garbanzo beans onto a baking sheet. Bake for one hour, stirring occasionally, until crisp and brown. Cool to room temperature before eating. (Store leftovers in an airtight container.)", "Nutrition Information", "(per serving)", "Calories: 176", "Fat: 3 g", "Saturated Fat: 0 g", "Cholesterol: 0 mg", "Sodium: 474 mg", "Carbohydrates: 30 g", "Sugar: 3 g", "Fiber: 7 g", "Protein: 9 g", "Calcium: 39 mg"]]
- "targetTitle": "The Anytime Snack You Won't Feel Guilty About Eating (VIDEO)"
- "targetDescription": "Whether you're between meals or craving a midnight snack, there's no need to feel guilty about getting your fill. This healthy, spiced snack is lower in ..."
- "targetKeywords": "Healthy Snacks, healthy snack ideas, healthy snack, Garbanzo Beans, chickpeas, Art Smith, art smith recipes, Harvest Day, #ownshow, ownshow, own-video, Oprah Winfrey Network"
- "targetMedia": ["http://i.huffpost.com/gadgets/slideshows/284066/slide_284066_2172537_free.jpg", "http://big.assets.huffingtonpost.com/FBLOGO_0.png", "http://s.m.huffpost.com/assets/Logo_HuffingtonPost-1f66cbef067771f8508930828f764bc8bf8c4b027e58e14e07d207a9f0ce1d31.svg", "<http://i.huffpost.com/gen/2472632/thumbs/o-SMOKY-PAPRIKA-BAKED-GARBANZO-BEANS-SNACK-570.jpg?6>", "<http://big.assets.huffingtonpost.com/TWITTERBIRD.png>", "<http://b.scorecardresearch.com/p?c1=2&c2=6723616&c3=&c4=&c5=&c6=&c15=&cj=1>", "http://s.m.huffpost.com/assets/callout_highlight_icon-930f97b89eddc7188ad1c1e0a96bc1359e94a46ea21239c3c5ca3f25c543eaeb.gif", "<http://s.m.huffpost.com/assets/snapchat-QR-56da75d622f38e76ff0d3dec97344808f972f51020435d37d55c6f0e2eaec780.png>", "<https://www.facebook.com/tr?id=10153394098876130&ev=PageView&noscript=1>", "<http://secure-us.imrworldwide.com/cgi-bin/m?ci=us-703240h&cg=0&cc=1&ts=noscript>", "<https://www.facebook.com/tr?id=1621685564716533&ev=PageView&noscript=1>", "<http://i.huffpost.com/gen/2472632/images/o-SMOKY-PAPRIKA-BAKED-GARBANZO-BEANS-SNACK-facebook.jpg>"]]
- "targetUrl": "<http://huff.to/1xNip52>"
- "provenance": {"source": "anonymized", "humanSpoiler": "Smoky Paprika-Baked Garbanzo Beans", "spoilerPublisher": "HuffPoSpoilers"}}
- "spoiler": ["Smoky Paprika-Baked Garbanzo Beans"]]
- "spoilerPositions": [[[1, 0], [1, 34]]]
- "tags": ["phrase"]}

Examples

- {"uuid": "c4e0edb3-16f0-49eb-b3a5-cdd24c546c74"}
- "postId": "1100188040002123_1111142778906649"
- "postText": ["You'll Never Believe What This Family Saw in the Sky Outside Their House in Finland."]
- "postPlatform": "Facebook"
- "targetParagraphs": ["\"It was cold and very foggy, the temperature was around -10 degrees Celsius,\" said Hänninen. \"When the clouds began to break, there were rainbow colours in the sky and a halo spanning 360 degrees! It was worth taking a picture or two.\\"", "If I ever stepped outside and saw this in my backyard, I might think the aliens were invading!"]
- "targetTitle": "You'll Never Believe What This Family Saw in the Sky Outside Their House in Finland."
- "targetDescription": null
- "targetKeywords": null
- "targetMedia": ["http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2015/01/subway-cars-145x100.png", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/04/water-blob.jpg", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/01/2-QJguJPe.jpg", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2015/01/viral-forest-logo-v6.png", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2015/01/mirror-shack-145x100.png", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/03/balloons-cov1.jpg", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/07/cemani.png", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/01/3-XmqQuvu.jpg", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/01/4-A0dsnoc.jpg", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/01/1-t4iFZ7u.jpg", "http://www.viralforest.com/wp-content/uploads/2014/01/3-XmqQuvu-600x350.jpg", "http://vrf.wpengine.netdna-cdn.com/wp-content/uploads/2014/03/borders.png"]]
- "targetUrl": "http://www.viralforest.com/youll-never-believe-what-this-family-saw-in-the-sky-outside-their-house-in-finland/"
- "provenance": {"source": "anonymized", "humanSpoiler": "\"Sun Halos\" Which are caused by ice crystals in cirrostratus clouds. #StopClickBait", "spoilerPublisher": "Stop Clickbait - Science"}}
- "spoiler": ["rainbow colours in the sky and a halo spanning 360 degrees"]
- "spoilerPositions": [[[0, 138], [0, 196]]], "tags": ["passage"]}

Examples

- {"uuid": "8841b82b-b3a7-4548-88d8-8507d4c9affb"}
- "postId": "565144999646724096"
- "postText": ["This guy pulled a really weird karaoke prank and no one noticed"]
- "postPlatform": "Twitter"
- "targetParagraphs": ["Anyone who likes karaoke has a go-to song. At first glance, it appears the guy in this video going viral on Reddit is going to sing Smash Mouth's \"All Star\" whether anyone wants to hear it or not — even if the sound system is blasting John Lennon's \"Imagine.\" The crowd does not seem to care though, as people are cheering and singing along to the chorus even as the words on the overhead screen are displaying the lyrics to the late Beatle's classic.", "Artists Who Have Never Won A Grammy VIEW GALLERY | 23 PHOTOS Queen John Rodgers—Redferns/Getty Images Queen", "John Rodgers—Redferns/Getty Images 1 of 23"]]
- "targetTitle": "This Guy Just Pulled a Really Weird Karaoke Prank and No One Noticed"
- "targetDescription": "Guy at karaoke bar sings All Star by Smash Mouth even though lyrics to Imagine by John Lennon are on the screen"
- "targetKeywords": "viral"
- "targetMedia": ["data:image/gif;base64,R0IGODlhAQABAAAAACH5BAEKAAEALAAAAAABAAEAAAICTAEAOw==", "https://timedotcom.files.wordpress.com/2015/02/karaoke-imagine-all-star-smash-mouth-john-lennon.jpg?w=720", "https://timedotcom.files.wordpress.com/2015/02/no-grammy-05.jpg?quality=85&w=838", "https://timedotcom.files.wordpress.com/2015/02/no-grammy-16.jpg?quality=85&w=838", "https://timedotcom.files.wordpress.com/2015/02/no-grammy-11.jpg?quality=85&w=838", "https://timedotcom.files.wordpress.com/2015/02/no-grammy-11.jpg?quality=85&w=405"]]
- "targetUrl": "http://ti.me/1EUHsHv"
- "provenance": {"source": "anonymized", "humanSpoiler": "Sang a different song's lyrics.", "spoilerPublisher": "SavedYouAClick"}}
- "spoiler": ["At first glance, it appears the guy in this video going viral on Reddit is going to sing Smash Mouth's \"All Star\" whether anyone wants to hear it or not — even if the sound system is blasting John Lennon's \"Imagine.\"]"]
- "spoilerPositions": [[[0, 43], [0, 259]]]
- "tags": ["multi"]}

Possible Approaches

1. Clickbait Spoiling via Simple Seq2seq Generation and Ensembling
 - This comprises two parts:
 - 1) simple seq2seq spoiler generation
 - 2) post-hoc model ensembling
 - Reference: [TohokuNLP at SemEval-2023 Task 5: Clickbait Spoiling via Simple Seq2seq](#)
2. Generating clickbait spoilers with an ensemble of large language models
 - This approach proposes an ensemble of fine-tuned large language models for clickbait spoiler generation
 - Reference: [Generating clickbait spoilers with an ensemble of large language models](#)

Clickbait Spoiling via Simple Seq2seq Generation and Ensembling

Seq2seq Spoiler Generation

1

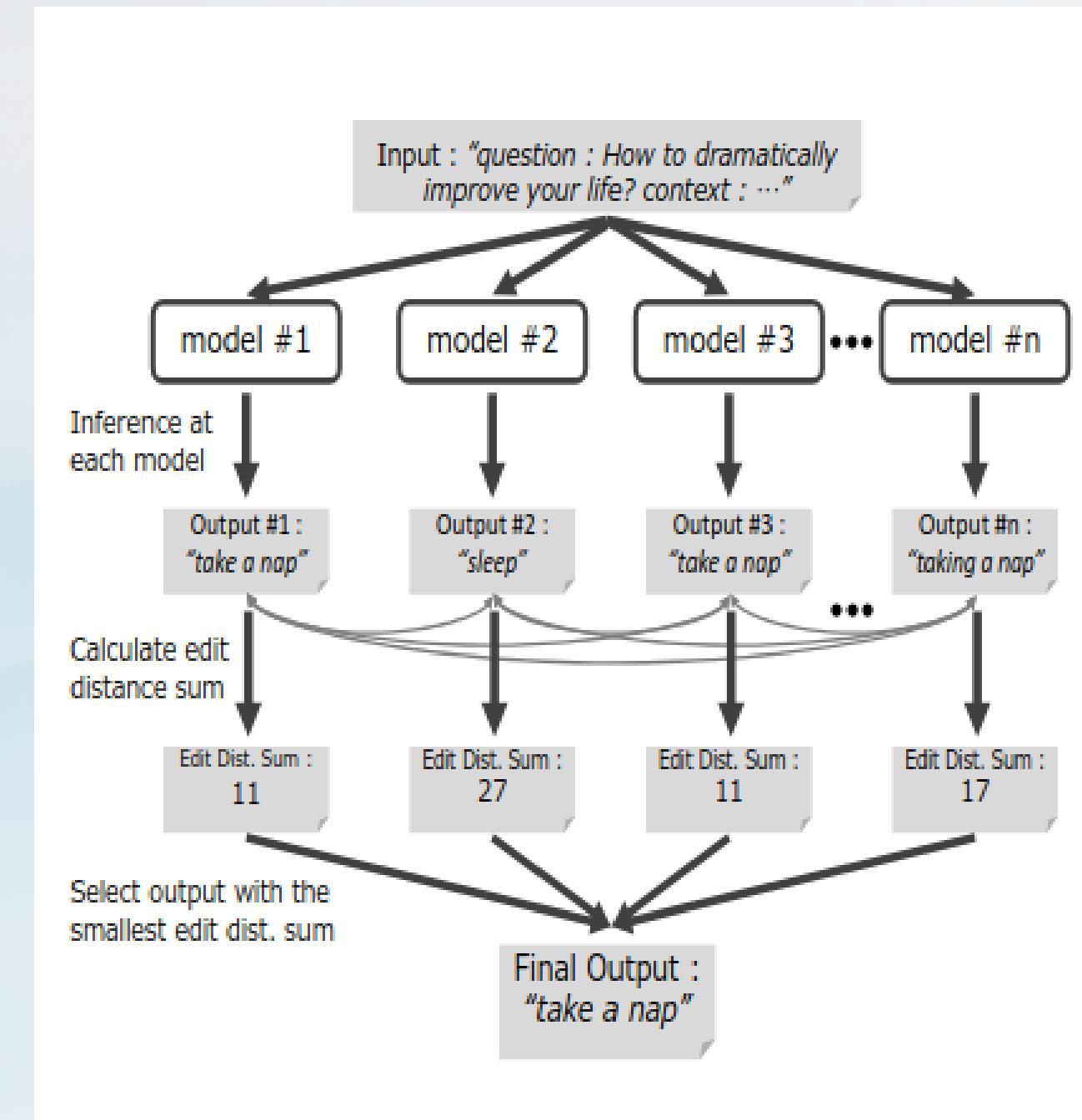
- Input format: "question : [postText] context : [targetTitle] - [targetParagraphs]"
- Output format: generates the sequence for the target spoilers directly

2

Post-Hoc Model Ensembling

- Takes predicted spoilers from various seed models as inputs, calculates the edit distances of each predicted spoiler
- Sums the computed distances over each model
- The final output is chosen by the spoiler with the smallest summed edit distances

Post-Hoc Model Ensembling



Generating clickbait spoilers with an ensemble of large language models

1

Converting clickbaits to questions

- QA datasets are often used as a part of LLMs' (pre)training sets
- To better exploit knowledge acquired by LLMs during pretraining, we convert each clickbait into a question

2

Generating spoilers with LLMs

- Using a set of different pretrained language models to produce a diversified set of spoiler candidates
- Each training example is converted to a textual form by filling in the hand-designed prompt template: "Below is a question paired with a context for which you should generate an answer. Write an answer with type t that appropriately completes the question.\n Question: q\n Context: a\n Answer: s\n"

Generating clickbait spoilers with an ensemble of large language models

3

Selecting the final spoiler with scoring model

- Construct the ranking of spoiler candidates and select the best candidate i.e. the spoiler at the top of the ranking
 - Pointwise approach:
 - Assigns score to each candidate, sorts, selects the top
 - During training, optimize the classical sum of squared errors
 - Pairwise approach:
 - Compares spoiler pairs, selects the most suitable
 - During training, optimise the standard cross-entropy loss

Evaluation Metrics

- BLEU (Bilingual Evaluation Understudy)
 - It calculates the overlap of n-grams (typically 1-4 grams) between the generated and reference texts, with a penalty for overly short sentences.
- BERTScore
 - Uses pre-trained BERT embeddings to capture token-level meaning and compares them between the two texts using cosine similarity.
- METEOR (Metric for Evaluation of Translation with Explicit Ordering)
 - It matches words using stemming and synonyms, and includes penalties for incorrect word ordering.

Plan & Proposed Timeline

- Week 1: Requirement Analysis and Setup
- Week 2 - 3: Baseline Model Implementation
- Week 4 - 5: Advanced Model Development
- Week 6: Error Analysis & Model Refinement
- Week 7: Evaluation and documentation
- Week 8: Adjustments and final submission

Thank You