

SOCIAL COMPUTING



Identifying Clickbaits and Unclickbaiting them

12th April 2021

Team 4

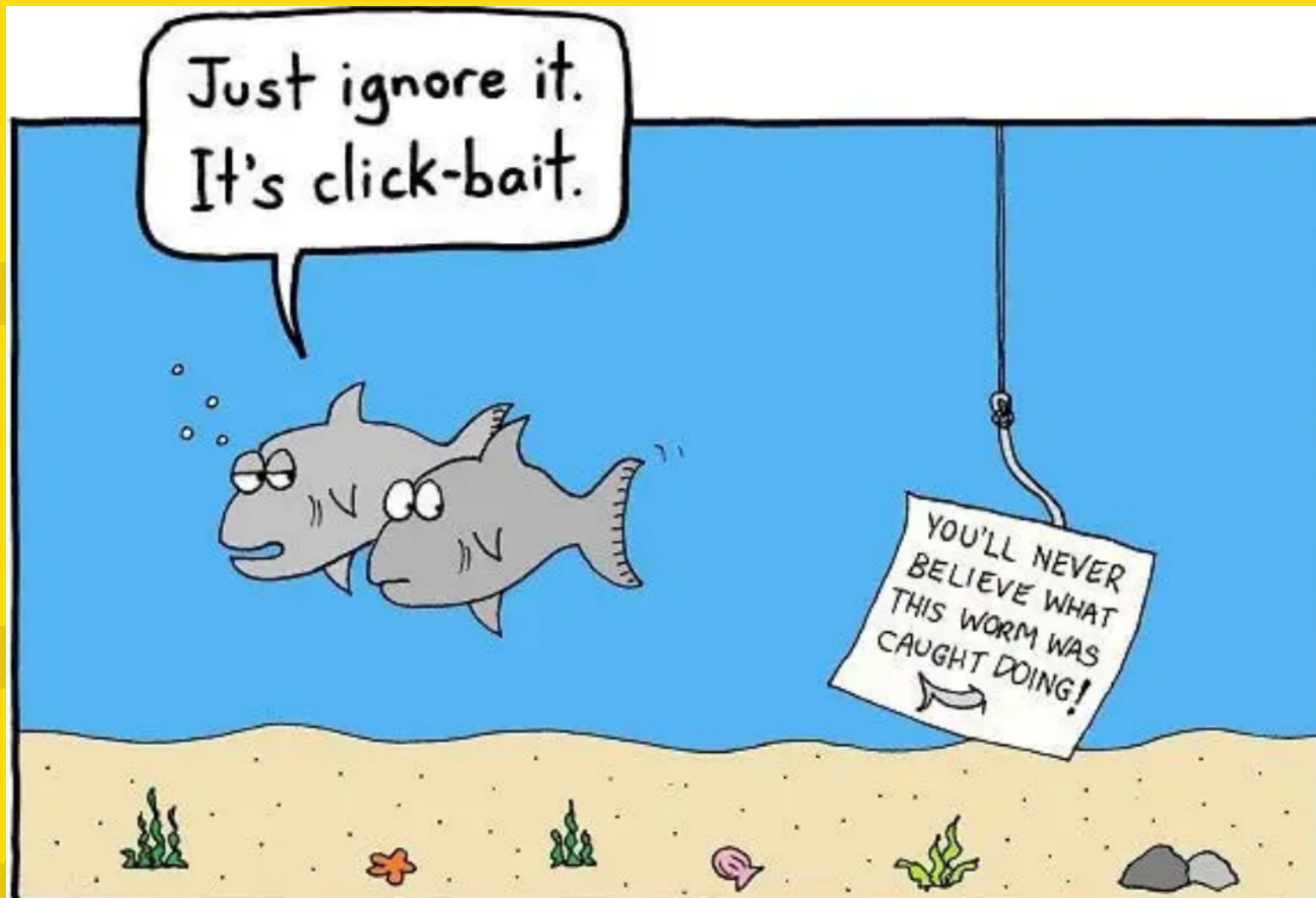


KEY DISCUSSION POINTS

- Introduction
- Problem Statement
- Approach Pipeline
- Clickbait Detection
- Un-clickbaiting

INTRODUCTION

What is a clickbait?



Clickbaits are catchy social posts or sensational headlines that attempt to lure readers to click.

These headlines exploit a user's **curiosity gap** and lure them to click on links that often disappoint them.

There are majorly two types of clickbaits **headline(text) based** and **thumbnail(image) based**.

INTRODUCTION

- "THESE WHITE TIGER CUBS ARE THE MOST BEAUTIFUL CREATURES YOU'LL SEE TODAY"
- "HERE'S WHAT REAL VEGANS ACTUALLY EAT"
- "BOW WOW HAD NO CLUE HOW TO KILL TIME DURING THE GRAMMYS AND IT WAS HILARIOUS"
- "WE KNOW WHO YOUR CELEBRITY HUSBAND SHOULD BE BASED ON ONE QUESTION"

A **Clickbait** typically has several of the following characteristics:

- An eye-catching and compelling headline
- Easily **skimmable**.
- **Funny** or memorable images/video.
- **Humorous** tone, or appeals strongly to a specific emotion.
- Intended to encourage social sharing

INTRODUCTION

The Benefits of Clickbait

- More Pageviews
- Greater **Potential** for Social Shares
- Increased Brand Awareness

The Dark Side of Clickbait

- Sensationalism Is Getting Old
- **Misleading** Clickbait Damages Brands and Erodes Trust
- Pageviews Aren't Even That Important



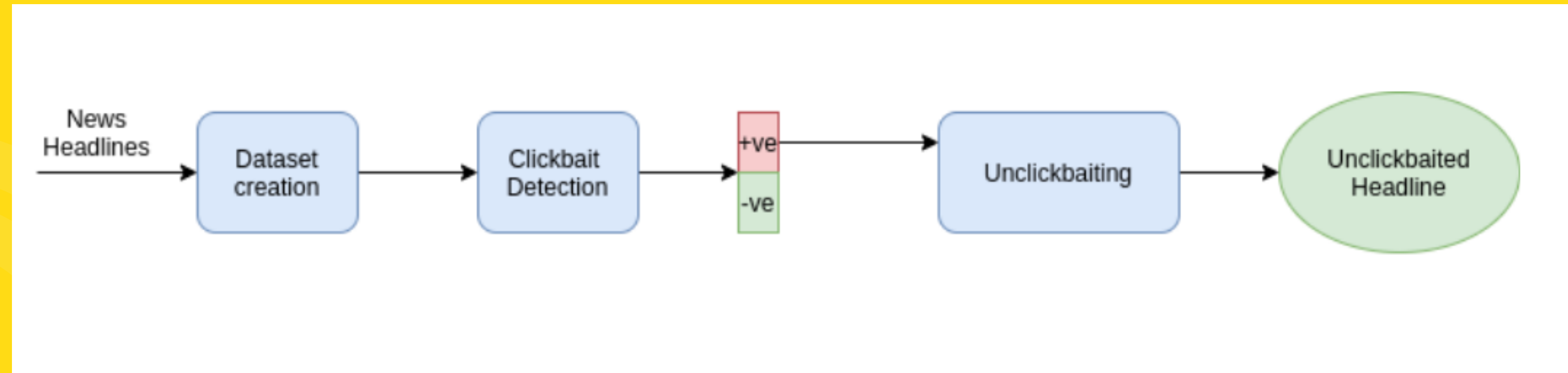
PROBLEM STATEMENT

**To identify clickbaits and
unclickbait it**

PROBLEM STATEMENT

- Clickbaits have been increasingly trending in most of the networking sites. So it is **important for us** to **detect and notify** users regarding **clickbaits** as it helps us to give a **good user experience** preventing users from spending extended time on social media and the **digital world**.
- It also shields the users from **inappropriate content** present in the articles which most of times is quite different from what was expected from the headline.
- An extension of this is to **unclickbait** the detected clickbaits for improved user experiences.

Approach Pipeline



CLICKBAIT DETECTION

Classifying a given headline as a clickbait
or not a clickbait

LITERATURE SURVEY

- Numerous works have been done on identifying clickbaits using NLP techniques.
- Most of the techniques involves training and validating a model on a given dataset comprising of **clickbaits and unclickbaits**.
- **Clickbait16K** was one of the first dataset which contained equal distribution of clickbaits and unclickbaits.
 - The data is collected from various news sites. The clickbait headlines are collected from sites such as '**BuzzFeed**', '**Upworthy**', '**ViralNova**', and '**Scoopwhoop**'.
 - The relevant or non-clickbait headlines are collected from many trustworthy news sites such as '**WikiNews**', '**New York Times**', '**The Guardian**', and '**The Hindu**'.

LITERATURE SURVEY

- Previous works done on this dataset include using traditional NLP features sentence structure, n-grams features, word patterns(cardinal numbers, unusual punctuation) and clickbait language(hyperbolic words, slangs etc) on top of machine learning models like SVM, Decision Tree and Random Forest to classify the clickbaits
- Later Deep Learning models, like RNNs(LSTMs, GRUs) were used to capture context for classifying.
- Recent studies have worked on predicting the strength of the headlines.

OUR APPROACH

The state-of-the-art transformer models improved the standard rnns with non-sequential nature, longer dependencies(using self attention) and use of positional embeddings.

To increase the amount dataset and update it to current style of **clickbaits**, we added clickbait headlines from the past recent years and for the **non-clickbaity** headlines were added from **abcnews** creating a dataset of **15 Lakhs** with equal proportion of both the types.

We **finetuned** the created dataset on pre-trained **BERT** and **RoBERTa** model as a **sentence classification task**.

The RoBERTa model owing to better performance we thought would serve as the "**groundtruth model**" to identify the clickbait for us to unclickbait.

RESULTS

Model	Accuracy	Precision	Recall	F1-score
Chakraborty et al. (2016) (SVM)	0.93	0.95	0.90	0.93
Chakraborty et al. (2016) (Decision Tree)	0.90	0.91	0.89	0.90
Chakraborty et al. (2016) (Random Forest)	0.92	0.94	0.91	0.92
BiLSTM(CE+WE)	0.98	0.98	0.98	0.98

Model	Accuracy	Precision	Recall	F1 macro
BERT	98.125	98.074	98.125	98.099
RoBERTa	98.406	98.425	98.392	98.405

LIMITATIONS

- On testing our model on daily headlines, we observed that the clickbaits from various websites including **buzzfeed and scoopwhoop** were **easily identified** by our model.
- When we tested on non-clickbait news headlines of various different sources(other than which we trained on) had inconsistencies because of varying linguistics styles. For Example Inshorts(whose headlines are attractive but not clickbaity).

UN-CLICKBAITING

Unclickbaiting the Clickbait

UNCLICKBAIT:

- Non-Persuasive.
- Unsensationalized/Non-exaggerated.
- Capture the content of the article.

STYLE TRANSFER:

- Language is rich and powerful. We can express the same meaning in different ways.

Ex: brief/verbose, Professional/unprofessional

- Text Style Transfer (TST) task, which aims to **change the stylistic properties** of the text while retaining its style-independent content.
- Style transfer is used for tasks like sentiment transfer, Story-level text generation, Shakespeare, negative to positive.

STYLE TRANSFER:

- Style transfer can be done using both supervised and non-supervised methods.
- **Supervised** : Supervised techniques require **parallel data** that contains a collection of original texts for the task and their corresponding expected set of outputs.
For example, in a style transfer task, parallel data contains a collection of original texts in Style 1 and their corresponding texts in Style 2.
- **Unsupervised** : These techniques do work with **Non-parallel data**. A non-parallel data is a data that contains a collection of original texts for our task and their corresponding expected set of outputs aren't available.

STYLE TRANSFER:

- Using massive amounts of parallel data has been essential in text generation tasks, such as machine translation and summarization. However, in many text generation problems, we can only assume access to non-parallel or monolingual data. Problems like style transfer are all instances of this family of tasks.
- Creating a parallel corpora is a tedious task. So, in our current task we will proceed with non-parallel data.
- To our knowledge, **no attempts have been made** so far to unclickbait a clickbaity title.

STYLE TRANSFER:

- Base architecture: (Style Transfer from Non-Parallel Text by Cross-Alignment, Shen et. al)

An Encoder :

+ Original sentence -> + content

+ style - style

A Decoder :

+ required style -> gives the desired output

+ content from encoder

- Drawback : Decoder will not only change the style-dependent content. It might also change the style-independent content.

STYLE TRANSFER:

- Improvised architecture: (STRUCTURED CONTENT PRESERVATION FOR UNSUPERVISED TEXT STYLE TRANSFER, Youzhi Tian et. al)
 - a) An Auto Encoder
 - b) Binary style classifier to generate sentence with the target style

DRAWBACKS OF STYLE TRANSFER IN UNCLICKABITING:

- Expectation was that it preserves the content of title
- Similar attempts in converting in generating clickbaity titles from non-clickbaity titles.
- But the other way around wasn't successful.
- Inference : Non-Clickbaity titles which are used as input in the previous model had a pattern, which was absent in our click-baits data.

Headline Generation using Summarization

Leveraging Summarization Models to generate unclickbaity headlines

DATASETS

- Extreme Summarization (XSum) Dataset
Contains around 200k articles along with 1 sentences summaries of that article.
- English Gigaword Corpus
Headline-generation on a corpus of article pairs from Gigaword consisting of around 4 million articles

T5

- Text-to-Text Transfer Transformer used for Abstractive Summarization.
- Its novelty lies in the ability to perform any sequence to sequence natural language processing task using just one trained model by the use of prompts.
- Use prompts such as "Summarize", "Translate to <x>", "Classify", etc. to do the respective tasks.
- We use a pretrained T5 model trained on CNN/Dailymail Dataset and we penalize larger summaries and also limit the maximum length to 40.
- Outputs non very human-like :(

ROBERTA2ROBERTA

- First proposed in "Leveraging Pre-trained Checkpoints for Sequence Generation Tasks", we use it for Abstractive Summarization
- RoBERTa initialized Seq2Seq model trained for Machine Translation, Text Summarization, Sentence Splitting, and Sentence Fusion
- The model is an encoder-decoder model that was initialized on the roberta-large checkpoints for both the encoder and decoder and fine-tuned on extreme summarization on the BBC XSum dataset, and another model fine-tuned on headline generation using the Gigaword dataset
- It seemed that model hallucinated a lot and produced headlines totally unrelated to the article.

BART

- "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension", proposed by Facebook.
- BART uses a standard seq2seq/machine translation architecture with a bidirectional encoder (like BERT) and a left-to-right decoder (like GPT).
- The idea around using BART is to train and store task specific models that generate very high quality data.
- It was a state-of-the-art abstractive summarization model before being outperformed by PEGASUS.
- Produced decent results.

PEGASUS

- PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization, proposed in 2019 by Google
- Pegasus' pretraining task is intentionally similar to summarization: important sentences are removed/masked from an input document and are generated together as one output sequence from the remaining sentences, similar to an extractive summary.
- Has state-of-the-art results on most (12) summarization datasets.
- Produced the best results among the 4 models according to human eye.

RESULTS

Gigaword

Model	ROUGE-1	ROUGE-2	ROUGE-L
ROBERTA2ROBERTA	38.62	19.78	35.94
BART	40.45	20.69	36.56
PEGASUS	39.12	19.86	36.24

xsum

Model	ROUGE-1	ROUGE-2	ROUGE-L
ROBERTA2ROBERTA	41.45	18.79	33.90
BART	45.14	22.27	37.25
PEGASUS	47.21	24.56	39.25

EXAMPLE OUTPUTS

Clickbait Title: "You Won't Believe this is the best way to save money"

Article: "Little known secrets time and headache."

Model Outputs:

T5: there are hundreds of sites offering free coupons from everything to grocery stores and shopping malls.

ROBERTA2ROBERTA: ["finding free coupons online", "Have you ever found a deal to find a discount on your supermarket or shop at a supermarket?"]

BART: ["how to find coupons online"]

PEGASUS: ['how to get free coupons on the internet _ step-by-step _ with bc-na-a & amp;']

EXAMPLE OUTPUTS

Clickbait Title: "Here's who the most hated man in the United States"

Article: "The most hated man way of life."

Model Outputs:

T5: by August 8, new regulations that began with Utah State Representative, may very well be the most

ROBERTA2ROBERTA: Do you know the man who brought the ban on electronic cigarettes in the United States ?

BART: ["utah 's most hated man in the united states"]

PEGASUS: ['this is an open letter to the most hated man in the united states.']

BROWSER EXTENSION

- Our Browser extension on a page ,
 - First check whether the given title to that page is clickbaity or not
 - If the title is clickbaity it generates the corresponding unclickbaity title.
- We used flask in the back-end, which is integrated with the neural models trained to predict the results. Further these results are used by browser extension.

17 Things Only Engineers With A Non-Engineering Job Will Relate

ADVERTISEMENT



GenerateNewTitle!

Newtitle :-Here's what it's like to go from an engineer to a non-engineering job.

There were zero protesters arrested the night before in Brooklyn Center.

17 April 2021, 16:31 • 6 min read

Apr 16, 6:04 PM

Apr 17, 6:58 AM



GenerateNewTitle!

No, Title is unclickbaity



Thanks!

Special Thanks to
Vijayasaradhi Sir and Nikhil
Sir for their feedback and
guidance