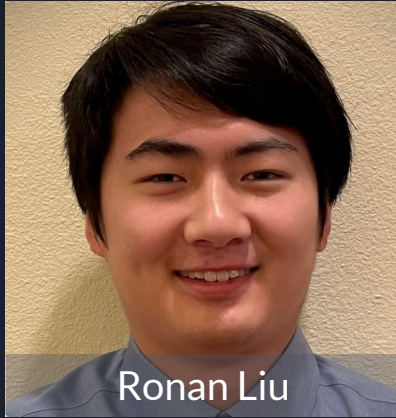# Satire Spotter

Barrett Hooper, Ronan Liu, Emerald Swei

The Satire Spotter Team

Ronan Liu

Software Engineer

Ronan aims to tackle the prevalence of misinformation being spread on social media in recent years.

Emerald Swei

Data Science Consultant

Emerald's primary ambition is to enhance online literacy and transparency, allowing users to make more informed decisions and react with context.
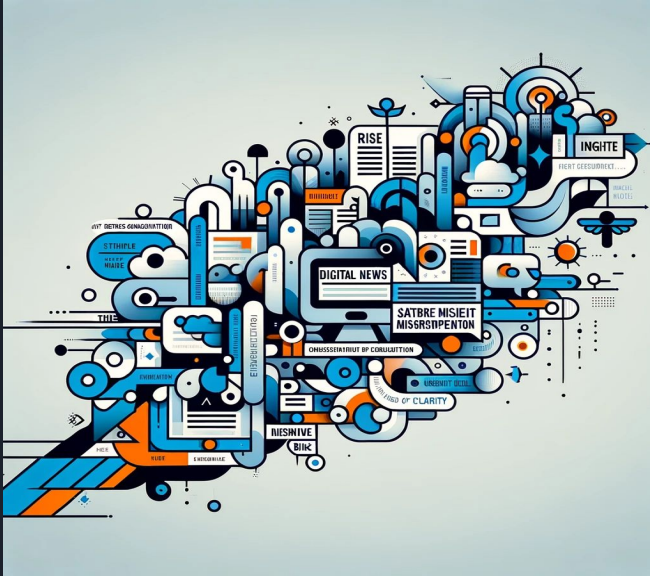
Barrett Hooper

Senior Financial Analyst

Barrett's interest is focused on mitigating the risks of satire misinterpretation, especially in a cross-cultural context.

Mission

Satire Spotter uses AI to detect satire in text and images, clarifying the media landscape and reducing misinformation to foster a better-informed, transparent digital society.

Problem Space and Impact

Rise of
Digital News
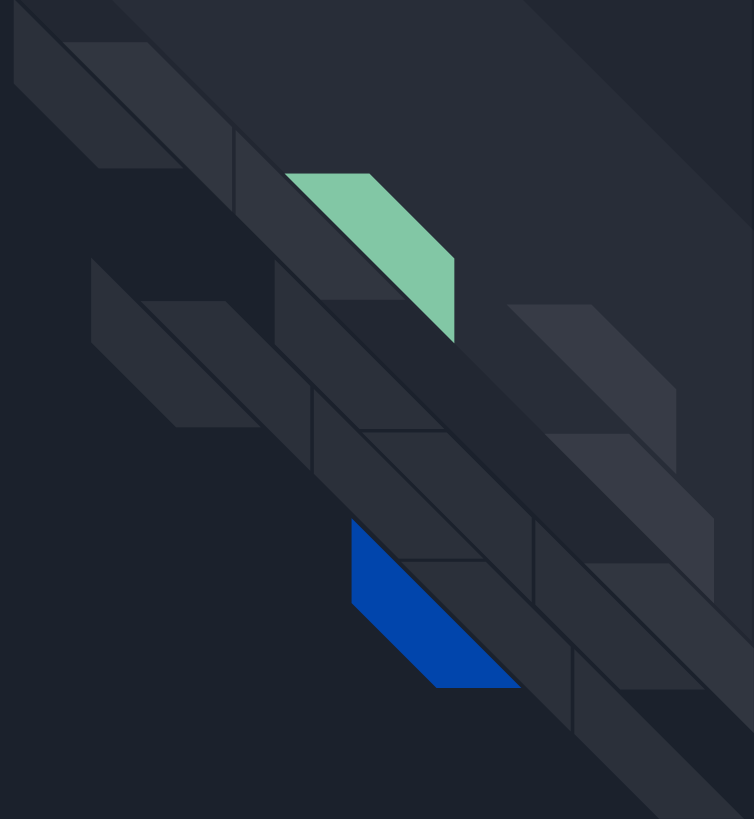Consumption

Risks of Satire
Misinterpretation

Urgent Need
for Clarity

Proposed Solution

Introduce a label at a granular, post level for all digital content to clarify the intent of each post and reduce confusion.

Technical Approach/Model Evaluation
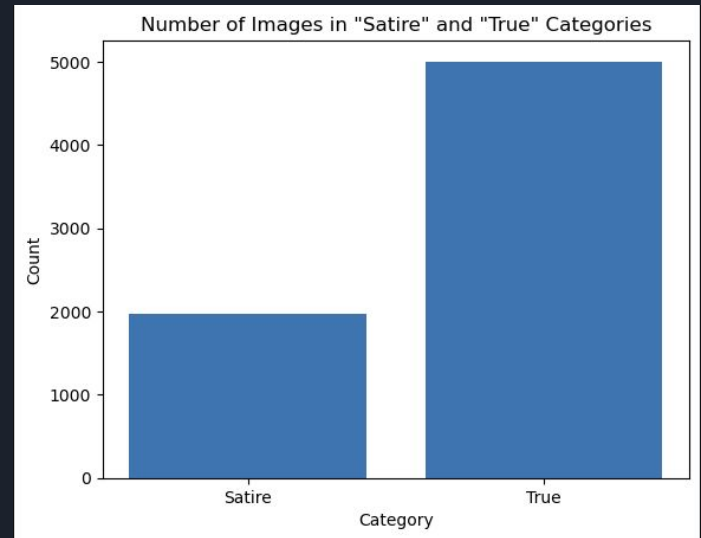
Datasets Used
(Approximate Numbers)

Text-only Model:

1. Satire News Detection Corpus
   230 Satire / 4,000 True
2. Crawled Current Onion Articles
   50 Satire
3. Princeton Senior Thesis Dataset
   13,000 Satire
4. Common Crawl News Dataset
   20,000 True

Images:

2,000 Satire / 5,000 True

Exploratory Data Analysis: Text Dataset

Word Clouds

Average Word Count Analysis

Top N Frequent Words

NER (Named Entity Recognition)

Bi-grams

Tri-grams

Sentiment Analysis

Average Sentence Length & Readability

Exploratory Data Analysis: Image Dataset

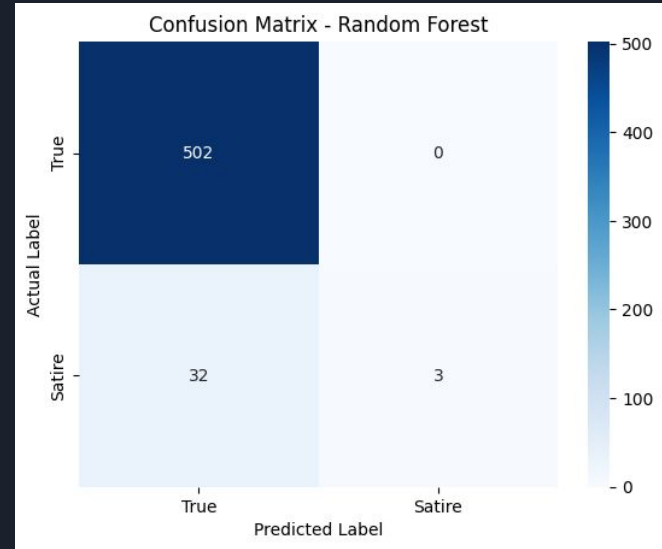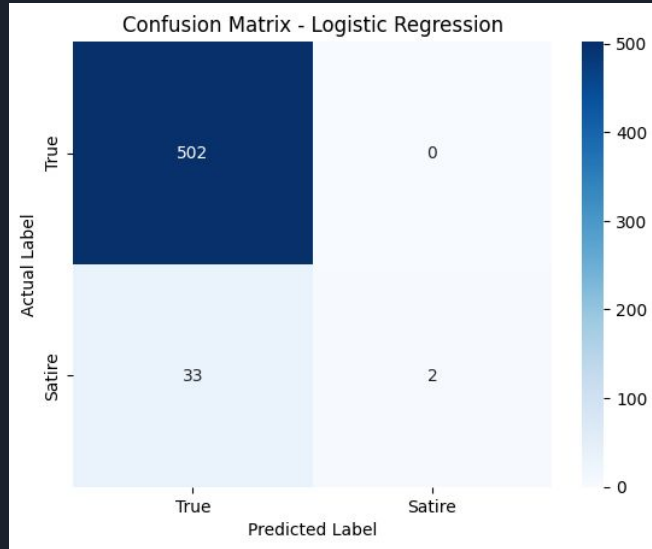Figure 1: Examples of satirical news images created by altering existing images.

Modeling Approach: Text

Logistic Regression 1.0 & Random Forest 2.0

Logistic Regression 2.0 & Random Forest 2.0

Long Short Term Memory (LSTM)

Logistic Regression 1.0 & Random Forest 1.0

Logistic Regression 2.0 & Random Forest 2.0

Hyperparameter Tuning

Regularization

Augmented Dataset

```
Classification Report - Logistic Regression:
                precision    recall  f1-score   support

        True       0.93      0.90      0.91      2662
      Satire       0.94      0.96      0.95      4501

    accuracy                           0.94      7163
   macro avg       0.93      0.93      0.93      7163
weighted avg       0.93      0.94      0.93      7163


Classification Report - Random Forest:
                precision    recall  f1-score   support

        True       0.88      0.84      0.86      2662
      Satire       0.91      0.94      0.92      4501

    accuracy                           0.90      7163
   macro avg       0.89      0.89      0.89      7163
weighted avg       0.90      0.90      0.90      7163
```

Important Features

Top influential tokens in Logistic Regression
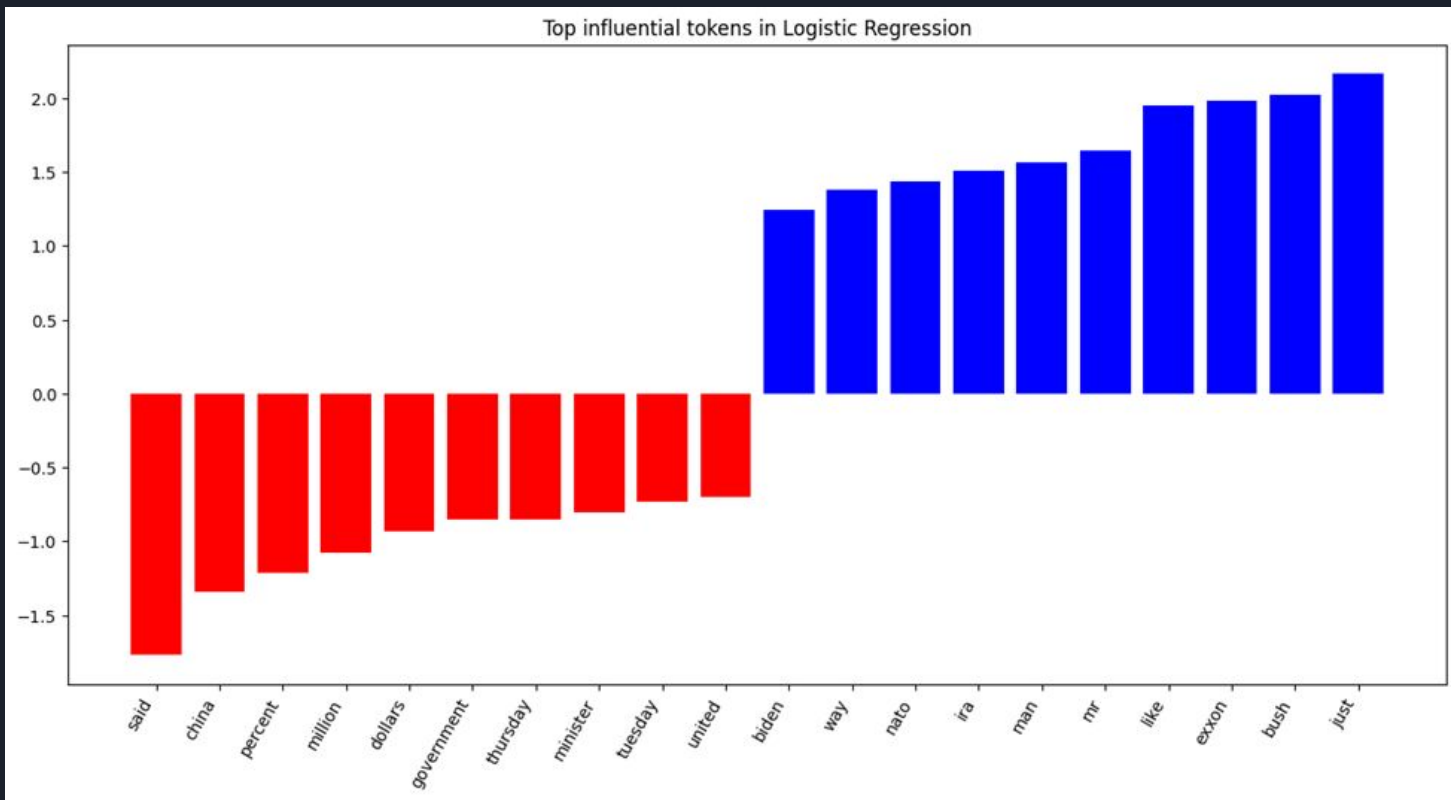
LSTM Deep Learning Model

```
224/224 [==============================] - 27s 116ms/step
              precision    recall  f1-score   support

        True       0.97      0.97      0.97      4501
      Satire       0.94      0.95      0.95      2662

    accuracy                           0.96      7163
   macro avg       0.96      0.96      0.96      7163
weighted avg       0.96      0.96      0.96      7163
```
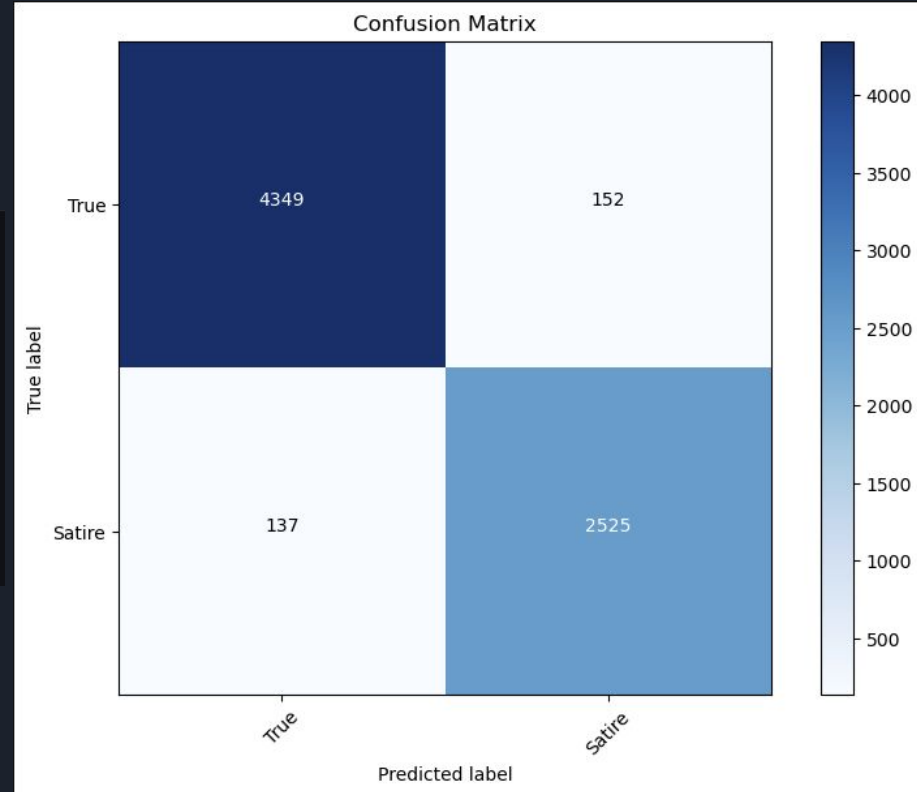

Confusion Matrix

Modeling Approach: Images

CNN	precision	recall	f1-score	support
True	0.64	0.83	0.72	134
Satire	0.77	0.55	0.64	141
accuracy			0.69	275
macro avg	0.71	0.69	0.68	275
weighted avg	0.71	0.69	0.68	275

SVM	precision	recall	f1-score	support
True	0.64	0.66	0.65	134
Satire	0.67	0.65	0.66	141
accuracy			0.65	275
macro avg	0.65	0.65	0.65	275
weighted avg	0.65	0.65	0.65	275

Pre-Trained Image Model


```
MobileNet V2

Classification Report:
              precision    recall  f1-score   support

           0       0.84      0.89      0.86       134
           1       0.89      0.84      0.86       141

    accuracy                           0.86       275
   macro avg       0.86      0.86      0.86       275
weighted avg       0.86      0.86      0.86       275
```
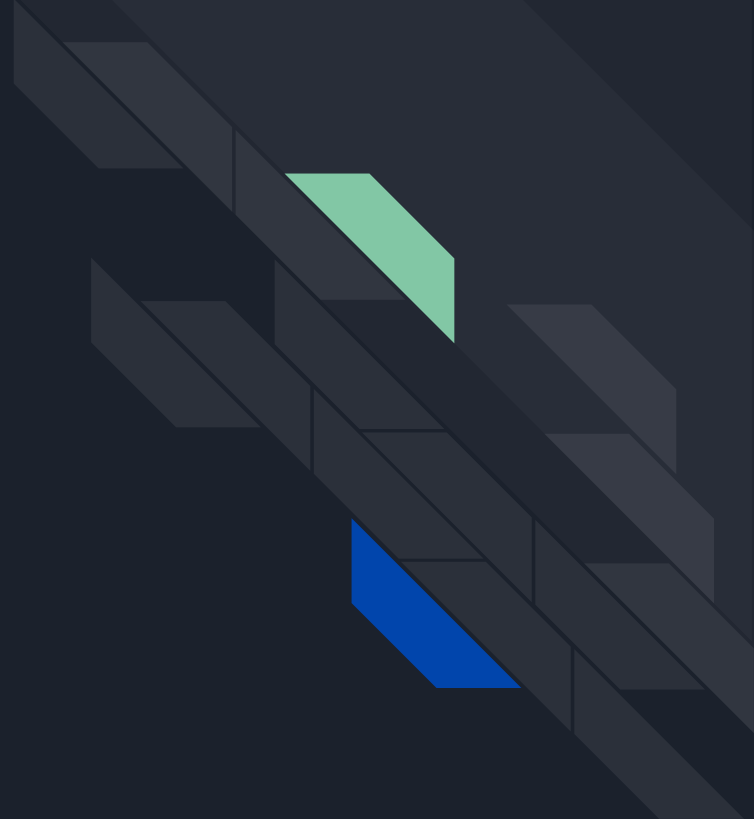
Final Findings

Comparison to Current Models

Text Model

Table II. Accuracy for all

Weighing Parameter	Accuracy
TF-IDF	83.4160 %
BIN	84.0833 %
BNS	89.9167 %
TF-BNS-IDF	92.7500 %

Table III. Comparative study

Weighing Parameter	Precision	Recall	F-score
TF-IDF	0.722	0.732	0.727
BIN	0.746	0.788	0.767
BNS	0.802	0.859	0.829
TF-BNS-IDF	0.826	0.870	0.840

Image Model

Type	Model	Accuracy	F1 score	AUC-ROC
	All regular news	60.00	—	50.00
Baselines	ResNet101	73.54	65.26	80.28
	$BERT_{BASE}$	91.33	88.64	96.77
	Simple fusion (average)	92.53	90.44	96.74
	Simple fusion (concatenation)	92.74	90.70	97.31
Proposed Models	ELA+CNN	44.20	51.86	44.61
	ViLBERT	**93.80**	**92.16**	**98.03**

Table 1: Model performance on satire detection dataset.

How Can Target Users Use Our Model?

Roadmap of Future Steps

While we are super happy about how far we've come, we recognize that we still have much opportunity for growth ahead of us.

Expand to new frontiers of mediums like video or audio

Work with social media platforms to integrate our models directly into their platforms

Collaborate with users to create a rating system and feedback loop

the ONION

Seaworld Employees Place Orcas In Plastic Bags Of Water While Cleaning Tank
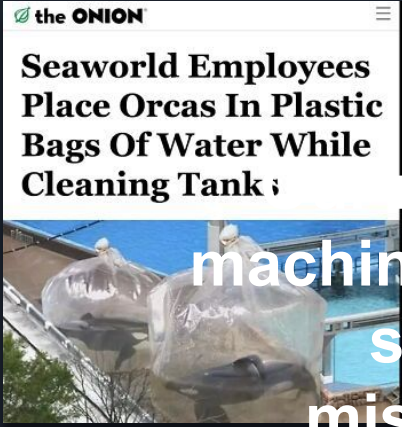
And this is why i refuse to go to Sea World

Chiquita ✓
@ChiquitaBrands

We've just overthrown the government of Brazil.

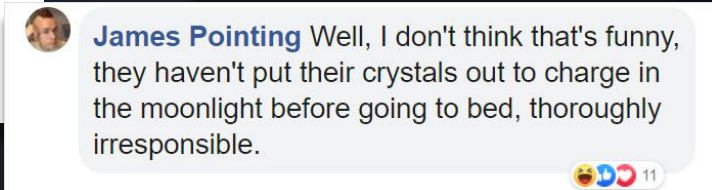11:52 AM · Nov 10, 2022 · Twitter for iPhone

Linked in

Yesterday I was walking to an interview. There was a starving dog on the road. I stopped to feed him & missed the interview. The next day I got a call asking to come in to do the interview. I was surprised, but I went. Then the interviewer came in. He was the dog

...leverage A... machin... cut out the s... mation from misinformation, and foster the development of a better... digit...

Jurassic Park Updates
@JurassicPark2go

If you are being eaten, please do not scream. You could potentially be spreading covid-19. Keeping visitors safe is our number 1 priority

12:45 PM · Jul 9, 2020

Funny Tweets · See more

Nestlé ✓ @Nestle...

We steal your water and sell it back to you lol

💬 224   ⟲ 8,682   ♥ 50.2K

Show this thread

James Pointing Well, I don't think that's funny, they haven't put their crystals out to charge in the moonlight before going to bed, thoroughly irresponsible.