Detecting Political Bias in News Articles Using BERT

Mehmet Inonu, Jian Wu

DATASCI 266, August 9, 2024

Motivations & Why It Matters

Existing Approaches

Berkeley

Political Bias Detections

Human Evaluations

- Allsides Media Bia Chart
- AD Fonte Media

Machine Learnings

- Support Vector Machines (Pang 2004)
- RNN/LSTM (Iyyer et. al 2014)
- BERT (Baly et. al 2020)

Sources: allsides.com

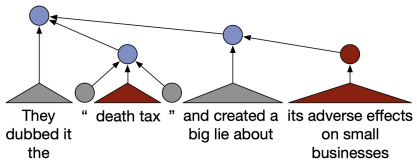


Figure 1: An example of compositionality in ideological bias detection (red → conservative, blue → liberal, gray → neutral) in which modifier phrases and punctuation cause polarity switches at higher levels of the parse tree.

Berkeley

Our Approach

Berkeley

BERT - A Refined Approach

Fine Tuning BERT Uncased Model by

- Comprehensive Data Cleaning and Preprocessing (Incomplete Data, Text Normalization, Source Anonymization)
- Experimenting with various Combination of Model Configurations and Hyperparameters (dropout rates, maximum token length, number of trainable layers, hidden layer dimensions, learning rates, and training epochs)
- Experiment with different model architectures (LSTM layer or CNN layer)
- Training and Test Set Variations (first paragraph of input text, one year data, 4 year data, removing top 5 sources data)
- Label Rebalancing

Berkeley

Results

Berkeley

Experiment Results I

Experiments	Train/Test Size	Accuracy	Weighted Average F1 Score
Naive Bayes Classification (Baseline Model)	20,114/5,029	69.00%	68.00%
BERT Uncased Fine-Tuning (CLS Token)	**20,114/5,029**	**91.00%**	**91.00%**
Baly et. al (2020)'s Fine-Tuned BERT (Random Split Dataset)	33,537/1,200	79.83%	80.19%

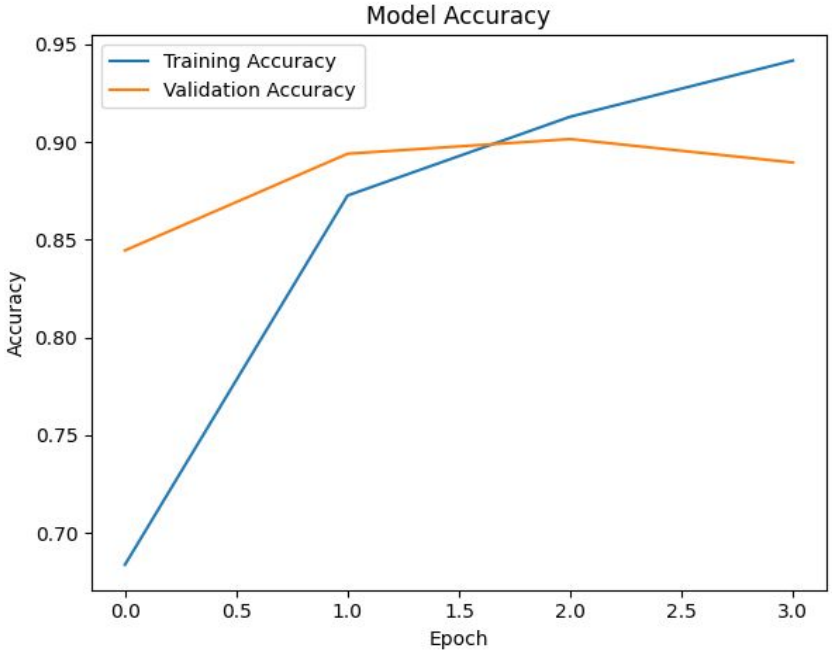
Berkeley

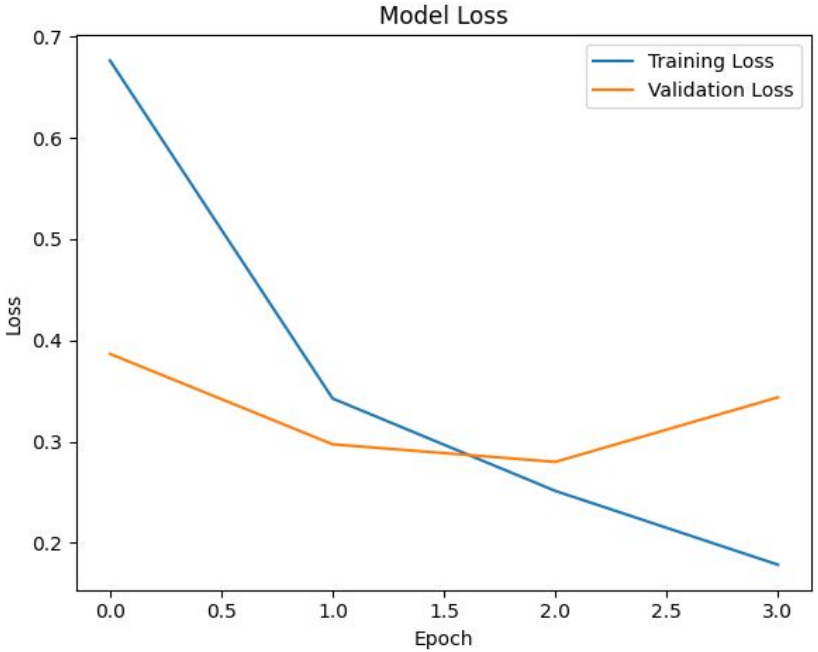
Experiment Results II

Experiments	Accuracy	Weighted Average F1 Score
BERT CLS Token (12 NN Layers)		
Fine-Tuned BERT	91.00%	91.00%
Fine-Tuned BERT with a LSTM Layer	**91.00%**	**91.00%**
Fine-Tuned BERT with a CNN Layer	83.00%	83.00%
BERT with Fine-Tune with the First Six Sentences		
Fine-Tuned BERT with First Six Sentences	**87.00%**	**87.00%**

Berkeley

Learning Curve





Berkeley

Additional Experiments I

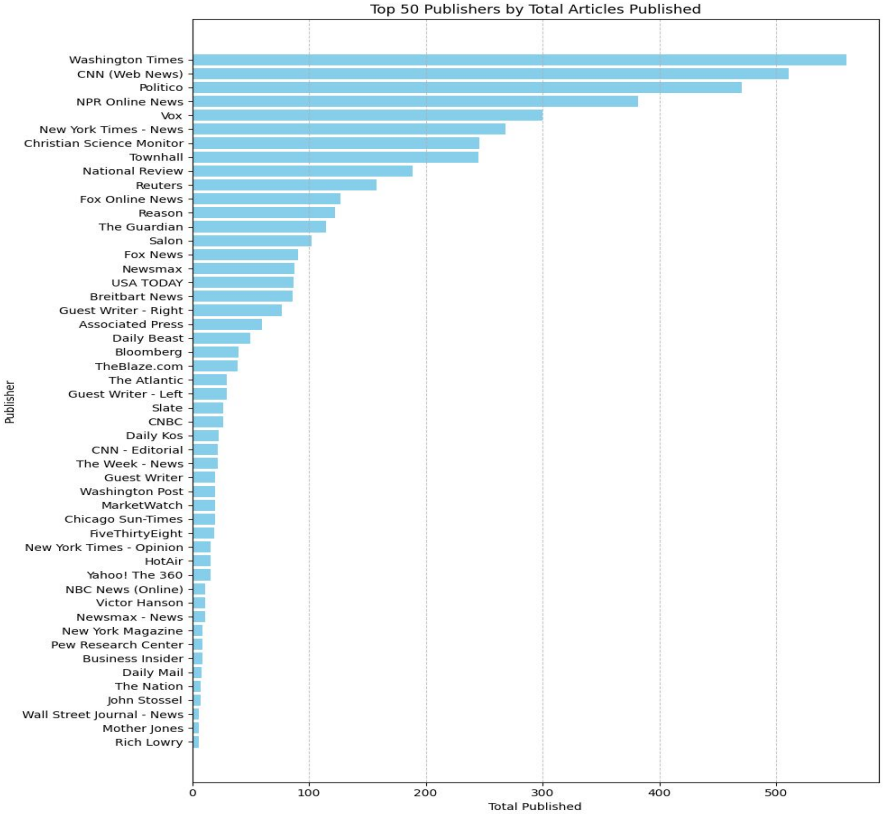
Experiments*	**Train/Test Split***	**Weighted Average F1 Score**
Fine-Tuned BERT with 2016 Data (Pooler)	2,893/4,407	79.00%
Fine-Tuned BERT **with 4-Year Data (CLS)**	**5,977/622**	**80.00%**
Fine-Tuned BERT without Top-5 Sources and Test on those Unseen Top-5 Sources' Articles	11,332/2,000 (3 batches)	40.00%

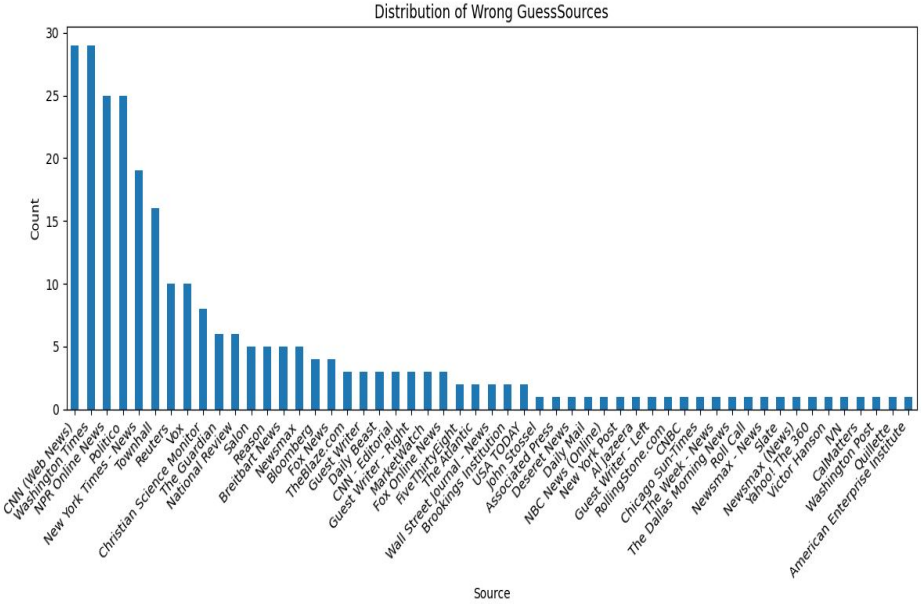
Berkeley

Error Analysis

Berkeley

Publication Volumes Proportioned to Erros





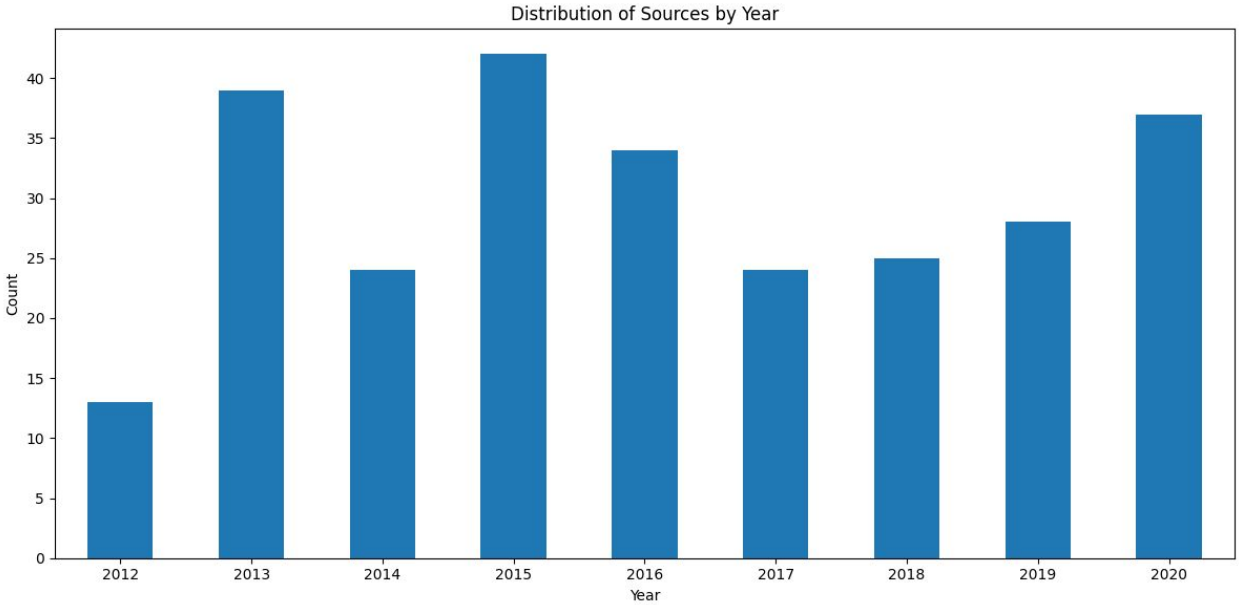
Berkeley

Additional Experiments II

Experiments	Train/Test Split*	Weighted Average F1 Score
Fine-Tuned BERT with Filtered Dataset (CLS)	**8,845/2,213**	**86.00%**
Fine-Tuned BERT with LSTM & Filtered Dataset (CLS)	8,845/2,213	85.00%
Fine-Tuned BERT without Top-5 Sources	**11,332/2,805**	**89.00%**

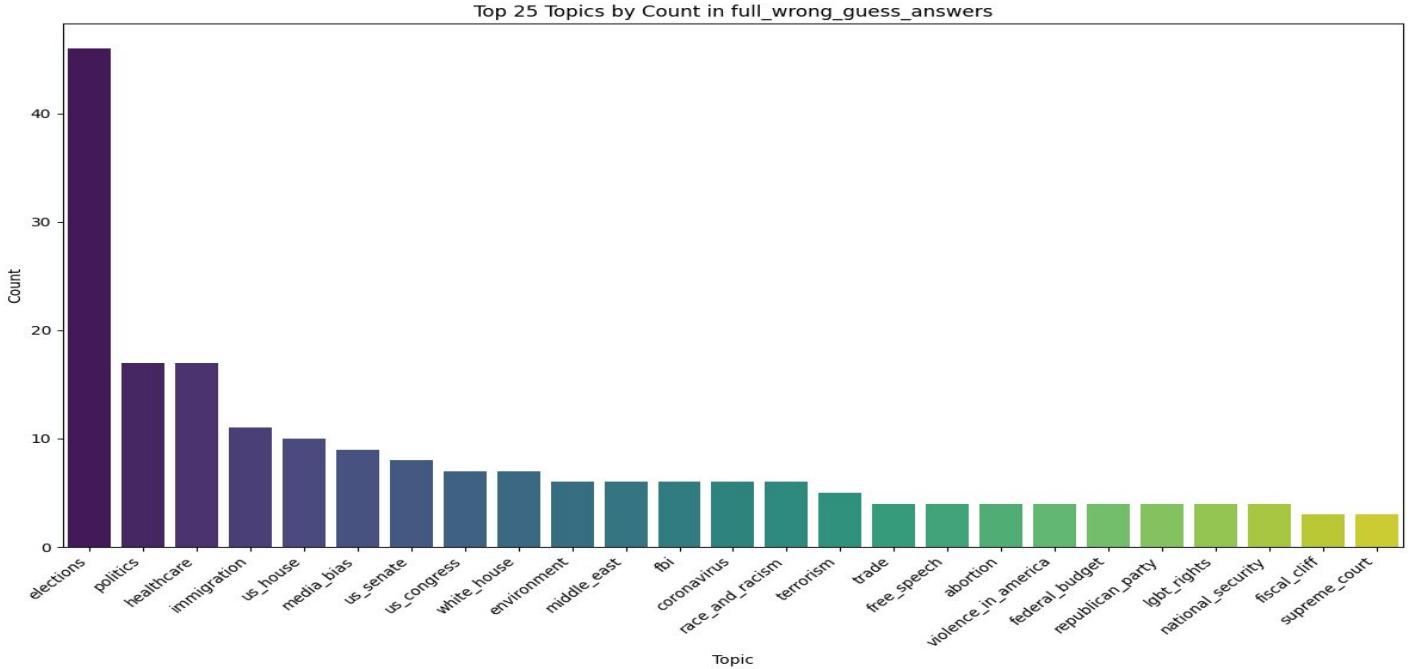
Berkeley

Publication Dates vs Erros



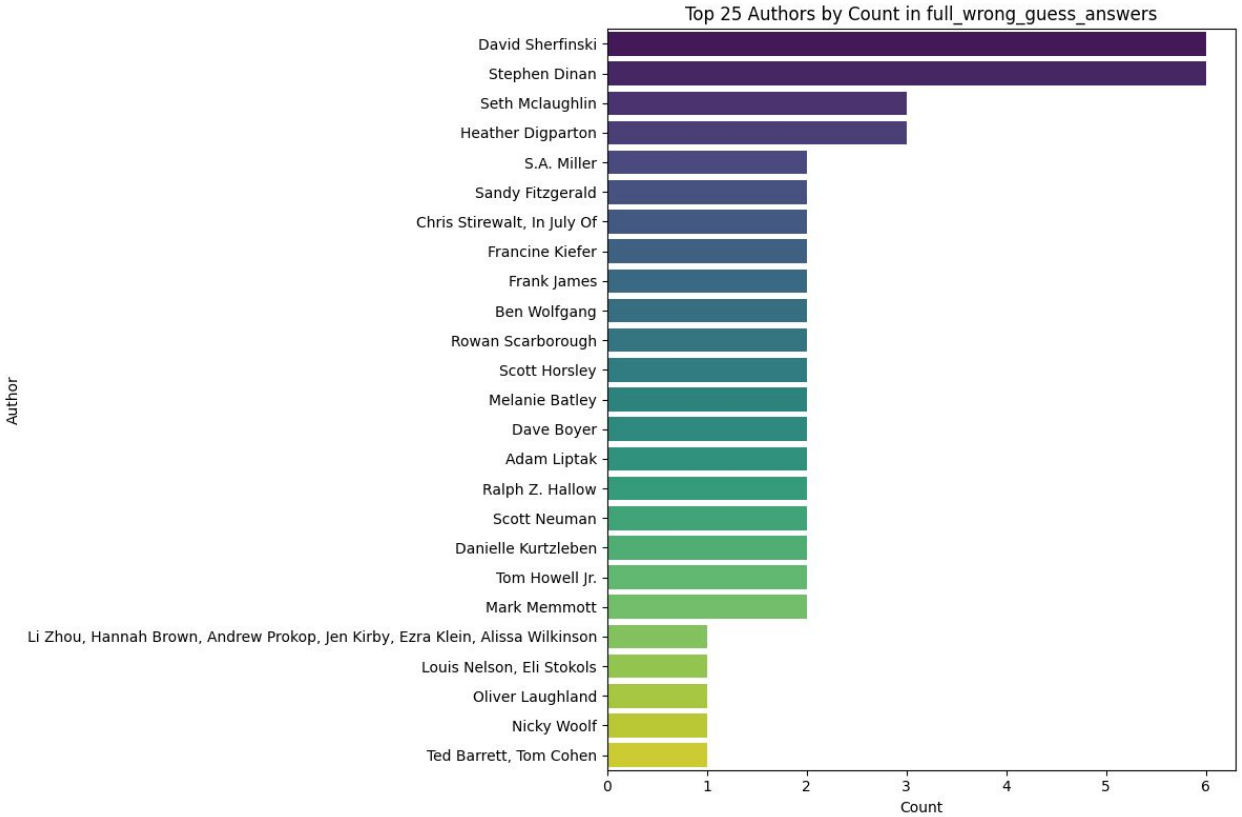
Berkeley

News Articles Topics vs Erros



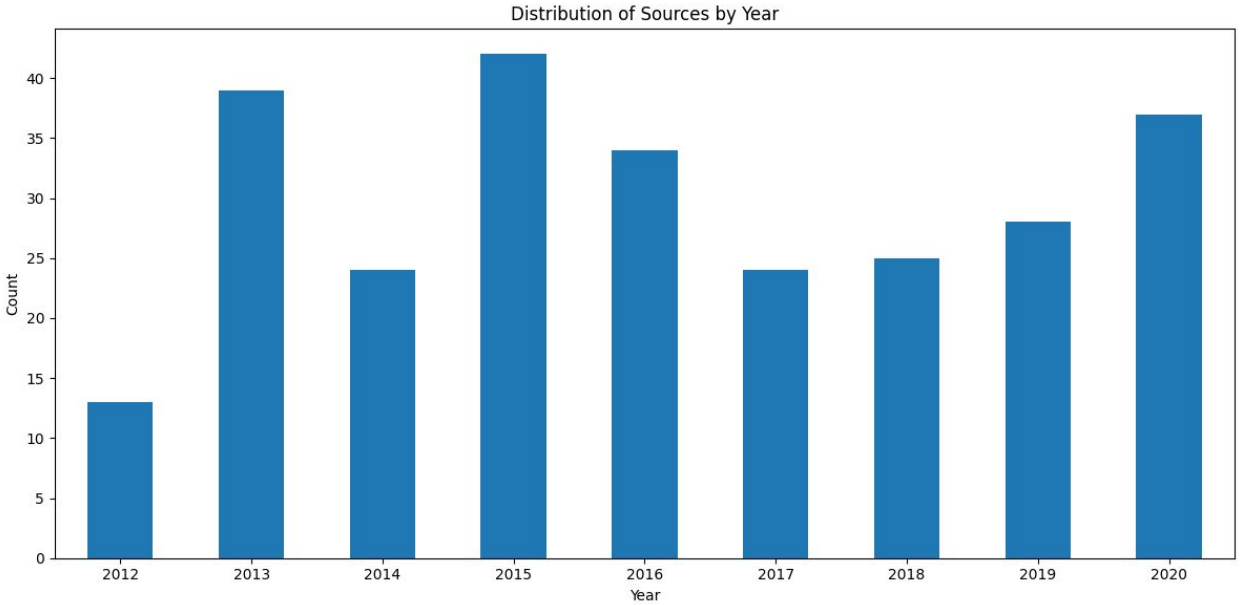
Berkeley

News Articles Authors vs Erros



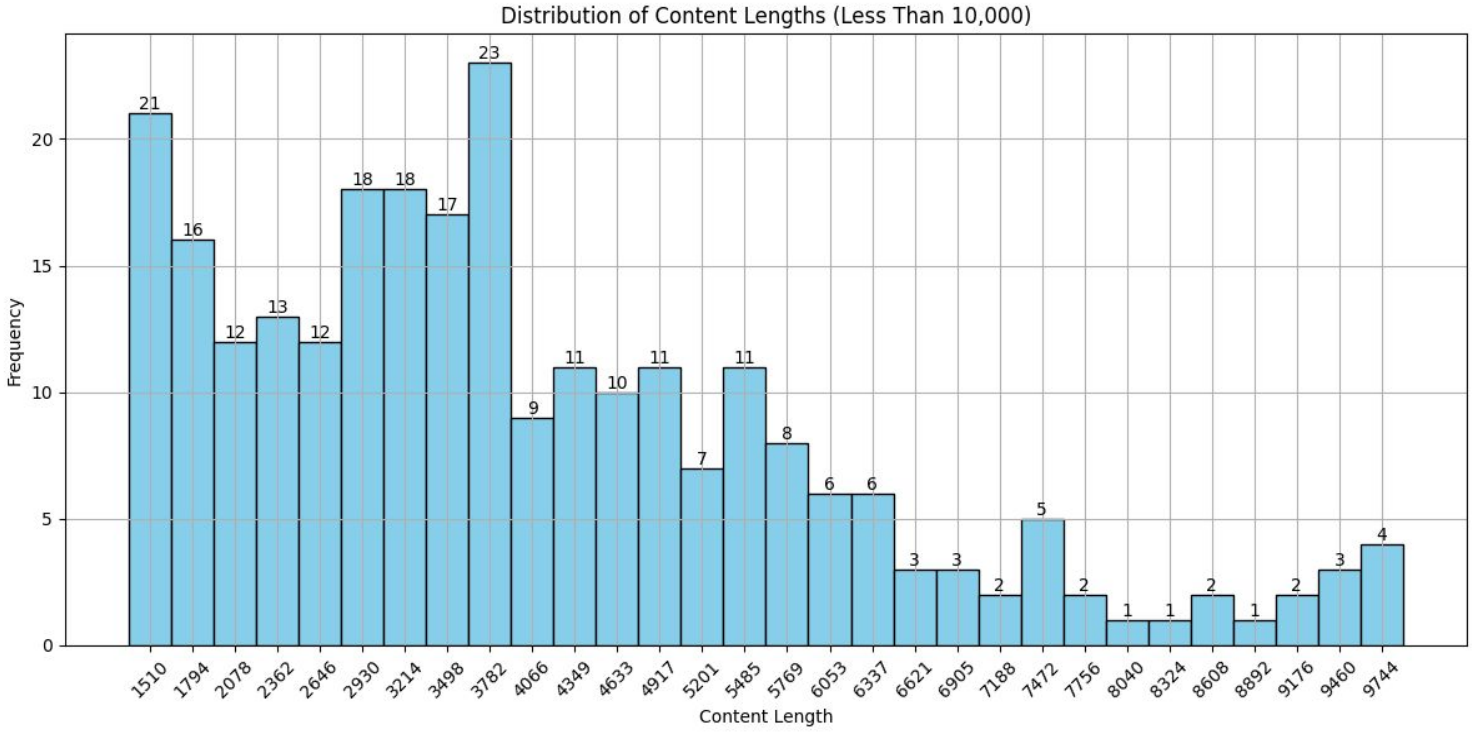
Berkeley

Publication Dates to Erros



Berkeley

Incorrectly Classified News Articles Length



Berkeley

Sample Error: 512 Limits the Full Understanding

Figure 10: Sample News Articles that the first portion is background information (True Label is Center; BERT model prediction is 2)

"Pesident barack obama ' s administration reportedly misled congress — american people — crucial element iran nuclear deal. a gopled investigation senate permanent subcommittee investigation revealed wednesday obama ' s treasury department secretly worked license allow iran access american financial system. a license, according treasury department, authorization [treasury] engage transaction otherwise would prohibited. " it ' s big claim. recall iran deal u lift certain sanction tehran exchange iranian curbing nuclear program. u didn ' t remove every penalty signing agreement 2015 — meaning iran ' s access u financial system still limited. but according new report, u treasury department issued license oman ' s bank muscat february 24, 2016. license would let iran change $ 5. 7 billion worth omani rial held bank euro — first changing u dollars. iran could change money dollar without license, since would violated sanction still place. however, two u bank refused help conversion iranianowned omani rial dollar — even obama official approached — reportedly citing severe hit reputations. all, knew would perceived circumventing iran nuclear deal aiding tehran. the associated press first report senate subcommittee ' s findings. sean kane, former sanction official obama ' s treasury department lawyer dechert llp, told treasury not practice allowing iran even indirect access u financial system. " he continued, reported appears limited authorization would allowed onetime conversion specific iranian reserve held abroad, something allowed continue accessing u financial system goingforward basis. " kane also noted treasury issue license based america ' s changing foreign policy needs. that ' s critic see it. obama effort boost iran ' s economy ran counter american interests, skirted law, " jonathan schanzer, iran expert conservative foundation defense democracy think tank, tweeted wednesday. did obama administration lie helping iran? what make revelation worse, republican lawmaker say, obama administration may best misled — worst, deliberately lied — granting iran limited access u financial system. the obama administration misled american people congress desperate get deal iran, " sen. rob portman (roh), subcommittee ' s chair, said statement. here ' s portman others make claim : wednesday ' s report, note statement top obama official saying type action wasn ' t possibility. iran continue denied access [us ' s] financial commercial market, " thentreasury secretary jack lew told senate foreign relation committee july 2015. and adam szubin, top treasury official obama, said senate banking committee later month : iranian bank able clear u dollar new york, hold correspondent account relationship u financial institutions, enter financing arrangement"

Berkeley

Proprietary Model’s Prediction

```
# ------------------------------------------------
# Input text:
# democrat congress seeking capitalize unrest teacher pay led classroom walkout handful states, offering electionyear plan roll back part last year ’ s gop tax cut boost fund schools,
#including higher wage educators. state local school district would get $ 50 billion decade raise teacher pay recruit educators. plan would provide another $ 50 billion school building resources,
#expand public school teachers ’ collective bargaining rights. the measure among series initiative democrat offering rally core voter ahead election november, every house seat ballot, along third
#senate seats. minority party chambers, democrat lack power bring plan vote. on monday, democrat focused controversy surrounding president donald trump allegation corruption administration, promising
#new ethic rule restriction campaign donation undisclosed contributors. to pay education plan, house minority leader nancy pelosi democratic leader want cancel new tax cut benefit wealthiest 1 percent a
#instead giving tax cut richest americans, give pay raise teacher country student depend succeed, ” senate majority leader chuck schumer said news conference. joining democrat leader american federation
#key donor party candidates. the recent teacher walkout occurred important election battleground state including arizona, gop senator jeff flake retiring, west virginia, incumbent democratic senator joe
#virginia ended nineday strike march 5 percent pay raise. state teacher demonstrated higher pay include oklahoma, kentucky colorado.


# True label:
# 1

# Predicted label:
# 2

# Chatgpt's Predicted Label (8k Tokens):
#TThe text reflects a left-leaning or progressive political ideology due to the focus on increasing funding for education, raising teacher pay, and opposing tax cuts for the wealthiest Americans.


# Claude 3.5 Sonet (200k Tokens))
#While the text does include some factual reporting of events, the overall framing and focus on Democratic proposals and critiques of Republican policies suggest a left-leaning bias. The language used,
#such as "Instead of giving tax cuts to the richest Americans," echoes typical left-wing rhetoric.
```

End
Any Questions?