Detection of Sensational Language Features in News Headlines

A Thesis Submitted
In Fulfilment of the Requirements
for the Degree of

MASTERS
In
Computational Linguistics

by
Po-Hsuan Chang
(33820070)

Under the Supervision of
Dr. AKSHI KUMAR
Goldsmiths, University of London

Goldsmiths
UNIVERSITY OF LONDON

To the

DEPARTMENT OF ENGLISH AND CREATIVE WRITING
DEPARTMENT OF COMPUTING
GOLDSMITHS, UNIVERSITY OF LONDON
New Cross, London SE14 6NW

September 2024

GOLDSMITHS, UNIVERSITY OF LONDON

CERTIFICATE

Date: __29/8/2024__

This is to certify that the work embodied in the thesis entitled ***"Detection of Sensational Language Features in News Headlines"*** done by Po-Hsuan Chang, 33820070 as a Post-graduate student in the Department of Computing, Goldsmiths University of London, UK is an authentic work carried out by him/her under my guidance.

This work is based on original research and the matter embodied in this research plan has not been submitted earlier for the award of any degree or diploma to the best of my knowledge and belief.

Supervisor

Dr. Akshi Kumar
Senior Lecturer and Director PGR,
Department. of Computing,
Goldsmiths, University of London
London, United Kingdom
akshi.kumar@gold.ac.uk

DECLARATION

I, Po-Hsuan Chang, Post-graduate student (33820070) in the Department of English and Creative Writing and the Department of Computing, hereby declared that the synopsis titled ***"Detection of Sensational Language Features in News Headlines"*** which is being submitted towards the fulfilment of the requirements for the degree of (Your Programme name) of Goldsmiths, University of London, United Kingdom is a record of bonafide research work carried out by me. I further declare that this work is based on original research and has not been submitted to any university or institution for any degree or diploma.

Po-Hsuan Chang
33820070
Dept. of English and Creative Writing
Dept. of Computing
Goldsmiths, University of London, United Kingdom
pchan003@gold.ac.uk / **changpohsuan@gmail.com**

Detection of Sensational Language Features in News Headlines

Abstract

This research acts as a pioneer of applying machine learning with natural language processing methods to identify a self-creating reannotation MIRUKU sensational news headline dataset, covering discussion of sensational language and identify important feature in news headline by SHAP analysis. All code in https://github.com/AD2000X/final_project.

Chapter 1: Introduction

1.1. Background of the Study

News:

The main source of information for modern people has transferred from broadcast, print, and television to the internet (Basera, 2015). News headline briefly illustrates the content of the news article, deliver the importance of the news article, and carry information for readers to decide whether to read the news content or not. Dor (2003) regards newspaper headlines are relevance optimizers: They are designed to optimize the relevance of their stories for their readers. Both above-the-line and below-the-line media aim to capture the attention of readers. The slang "if it bleeds, it leads" reveals that the strategy taken in the newsroom has never been changed. Their goal is not only to attract viewers' attention but also to keep their attention as long as possible. In this "15 minutes of fame" era, chasing higher CTR（Click Through Rate）is held up as a model for online media. Sales volume and viewership ratings are separately considered as the successful criteria of print media publishment and mass media.

Language:

Davis and McLeod (2003) indicates that the evolution of language is to promote the graduate complex social interaction and information exchange, making human individuals able to effectively deliver messages in the group. The point of view emphasized the importance of language in human society rather than just an intellectual tool. Uribe and Gunter (2007) argues that viewers can have emotional reactions to the news. Sensationalism in the news elicits emotion and sympathy from viewers. In news reports, sensationalism features contain features that can awaken audiences' emotional reactions and stimulate or ignite psychological responses (Uribe and Gunter, 2007).

Sensational:

Sensationalism is a characteristic that is emphasized in the process of news packaging, focus on elements that could stimulate human sensory device (Uribe and Gunter, 2007). Indicates that verbal measures of emotionality demonstrate remarkable correlations with physiological measures of emotional responses (Uribe and Gunter, 2007). Using sensational language in news affects the audience's attitude and cognitive process, such as formation and retention (Astari et al., 2023). The attention-drawing effect of sensational events can spread rapidly and widely, it also risks changing the fictional message into a fact during the propagation. The viral media can thrive in any context, as long as the information is sensational enough. The phenomenon can initiate with a conscious share or edit behaviour, evolution, and ends in everyone's information-gathering channels (Vista, 2014).

Evolution:

News sensationalism is not only often focused on the capability of attracting viewers' attention or stimulating physiological reactions, but also related to human evolution (Vettehen and Peeters, 2008). From the evolution perspective, the human brain forms a tendency that attends to the information that affects endurance and reproduction. The tendency makes humans naturally stay highly vigilant to potential threaten (Vettehen and Peeters, 2008). We further examine the possible reasons for sensational feelings in human emotions. From an evolutionary perspective, the goal of biological instinct is to reproduce the next generation. Vettehen and Peeters (2008) explain that sensationalism is not only often focused on the capability of attracting viewers' attention or stimulating physiological reactions, but also related to human evolution. The human brain forms a tendency that attend to the information that affects endurance and reproduction. This tendency makes humans naturally stay highly vigilant to potential threats and explains not only why sensational news can attract viewers, but also illustrates news reporters and audiences could appeal to deviant or negative news events. Therefore, the sensational news report can evoke people's adaptive pattern to pay attention to the information that could attract humans (Davis and McLeod, 2003). This evolutionary feature explains not only why sensational news can attract viewers, but also illustrates news reporters and audiences could appeal to deviant or negative news events (Vettehen and Peeters, 2008).

To capture readers' attention, people utilize different ways to construct an attractive headline, such as rhetoric skills, stimulating curiosity, and sentiment arousal to deliver the message to readers. Since the sense of elements of sensational are embedded in our genes to adapt to human evolution, it is necessary to identify the linguistic features that could arouse human attention. The approach and method of this research on news headlines are not only suitable for the media field but also act as an inspiration for different fields that language could highly affect human's biological decision system.

1.2. Problem Statement

Most of the research on sensational language is conducted in corpus analysis, and fewer are conducted by NLP (Natural Language Processing) programming techniques. The advantage of using NLP with machine learning is to develop a system that can effectively identify sensational language and understand what kind of elements account for a larger proportion of sensational language.

1.3. Purpose of the Study

Linguistics transition:

Categorizing news is a common practical, however, we can always find a news headline that contains multiple categorized elements. For example, (6) A pedophile monk who paid a schoolboy 50p each time he sexually abused him is locked up for five years (28 Jan. 1) (Molek-Kozakowska, 2013), this news title contains finance, sex, children, morals, and crime. Brown et al. (2018) suggests that sensationalism should be examined under digital circumstances, and treat sensationalism as a stylistic approach that combines classification and formation. Our goal in this research is to examine the linguistic features in news headlines.

1.4. Research Questions or Hypotheses

RQ 1. Can we utilize a single linguistic feature to effectively and stability identify the sensational language in news headlines?

RQ 2. What algorithm demonstrates robust performance in identifying sensational language within news headlines?

RQ 3. Which features most effectively identify sensational language in news headlines among selected features?

1.5. Scope and Delimitations

This research covers different linguistic features, including the Number of Words in the Headline; the Number of stop words in the headlines; The ratio of the number of stop words to the number of content words; TF-IDF with stop words; TF-IDF without stop words; Superlative word list; Syntactic 4-grams; Sentence Subjectivity and Objectivity Evaluation; Sentiment Analysis; Informality (Flesch-Kincaid Readability); Elongated Words; and Punctuation such as Currency symbols, Exclamation marks, Question marks, Ellipsis, Emphasis marks, Multiple exclamation marks, Single quotes, Double quotes and Contracted word forms. We use built-in stop words list of spaCy, meaning that different stop words lists could result in different consequences. We dropped features that are on our list, such as POS(part-of-speech) tag and POS N-gram. We calculate an optimal threshold of XGBoost but we didn't adjust the threshold in our follow experiment. We chose Cuckoo Search as our feature extraction method and chose Random Search as our model training method, thus the result could not be the same when anyone tried to reproduce the experiment. This research does not include a clustering algorithm for determining the topic of news headlines.

1.6. Definition of sensational

We define "Sensational Language" as follows:

Sensational language refers to a form of expression that quickly elicits strong emotional arousal and instant interest from the audience by appealing to their curiosity, emotion, or bias to capture attention and resonate, using vivid, exaggerated, or dramatic words, shocking details, or provocative content (Vista, 2014; Vettehen and Kleemans, 2017; Uribe and Gunter, 2007; Bell, 2015; Molek-Kozakowska, 2013; Brown et al., 2018).

Chapter 2: Literature Review

2.1. Overview of the Literature

Table 1

#	Type	Title	Publication	Contribution	Dataset	Implement	Feature	Metrics
1	clickbait	8 amazing secrets for getting more clicks': detecting clickbaits in news streams using article	Brown et al., 2018	Proposed a first automated detection method based on informality that is effective in identify clickbait titles.	4,073 news webpages on Yahoo homepage, sourced from various news sites such as	Lingua-EN-Tagger module of CPAN, Gradient Boosted Decision Trees	superlative (adjectives and adverbs), quotes, exclamations, use of upper case letters, asking questions, etc.; title-body	Precision, Recall, F-1 score, True Positive Rate (TPR), Feature Importance

informality	Huffington Post, New York Times, CBS, Associated Press, and Forbes.	Similarity (tf-idf); unigrams and bigrams form title and body of a page as features; subjectivity analysis; sentiment analysis; Informality; Coleman-Liau score (CLScore); RIX and LIX indices; Formality measure;

							web slang; swear words; pronouns and demonstratives; words containing repeated characters; Forward-reference; URL	
2	clickbait	A deep model based on Lure and Similarity for Adaptive Clickbait Detection	Zheng, Yu and Wu, 2021	Propose an adaptive deep learning model(Lure and Similarity for Adaptive Clickbait Detection (LSACD)) based on the degree of bait and the	Webis Clickbait Corpus 2017	LogisticRegression, RandomForestClassifier, DecisionTreeClassifier, GaussianNB, SVM; ConcatenatedNNArchite	Glove 300	Accuracy F1

				similarity between the title and content that can detect clickbait effectively.		cture, HybridModel, ZingelClickbaitDetector, RCNN + GRU, LSDA.		
3	clickbait	A Novel Contrastive Learning Method for Clickbait Detection on RoCliCo: A Romanian Clickbait Corpus	Broscoand Ionescu, 2023	Propose a contrastive learning model to detect Romania clickbait title and create the first Romania clikcbait corpus (RoCliCo).	RoCliCo (Romanian Clickbait Corpus).8,313 news samples which are manually annota	RF and SVM based on handcrafted features, Random Forest, SVM; BiLSTM network, Fine-tuned Ro-BERT	part-of-speech, tagging, scores (CLScore, LIX, and RIX), punctuation patterns	Precision Recall F1 Score

of News Articles.	ted with clickbait and non-clickbait labels. Romanian news websites (Cancan, Digi24, Libertatea, ProTV, WowBiz, Viva). 4,593

					non-clickbait and 3,720 clickbait.			
4	clickbait	BaitBuster-Bangla: A comprehensive dataset for clickbait detection in Bangla with multi-feature and multi-	Imran, Shovon and Mridha, 2024	Construct a multi-modal Bengali YouTube clickbait dataset (Mendeley Data clickbait dataset).	BaitBuster-Bangla	MiniLM-L12-v2, mpnet-base-v2, xlm-r-multilingual-v1	Metadata Features, Primary Content Features, auto_labeled, Human Annotation: human_labeled, AI-based Labeling: ai_labeled	Overall Accuracy (ACC), F1 Macro, F1 Micro, and Kappa scores

		modal analysis						
5	clickbait	BanglaClickBERT: Bangla Clickbait Detection from News Headlines using Domain Adaptive BanglaBERT and MLP Techniques	Joy et al., 2023	Build a large multi-modal Bengali clickbait detection dataset that provides debiased and human-annotated for low-resource languages and supports cross-language clickbait detection.	Annotated Dataset: Bangla Bait Unannotated Dataset	Logistic Regression, Random Forest; Masked Language Model (MLM), Ensemble of Convolutional neural network + Gated recurrent unit, Bengali GloVe Pretrained Word Vectors; LSTM, BiLSTM,	TF-IDF, n-grams(1-5), Bangla pretrained word embeddings, punctuation frequency, normalized Parts-of-Speech frequency, Abugida Normalizer and Parser for Unicode Texts (bnunicod	Precision Recall F1 Score

						BanglaBERT, XLM-RoBERTa	enormalizer), t-SNE	
6	clickbait	Believe it or not! Identifying bizarre news in online news media	Indurthi et al., 2018	We collect 23754 news headlines as bizarre news, sourced from news portals and channels exclusively catering to bizarre news. We develop and evaluate the first bizarre and unusual news items detection model.	Bizarre News Dataset (Weird), Conventional News Dataset	Multi-Layered Perceptron, Support Vector Machine (SVM) RBF kernel, Random Forest, Logistic Regression, XGBoost, Convolutional Neural Network (CNN),	Sentence Structure and Punctuation: Length of the News Headline, Stop words, Quotations using Colons, Quoted Content, Ellipses; Linguistic Patterns: Frequency of Popular Subjects,	Precision Recall F1 Score

Pronouns

and

Possessive

s,

Acronyms

and

Abbreviati

ons;

Named

Entities:

Persons,

Locations

and

Organizati

ons (PLO),

Country

Names,

Animals

and Plants,

Human

and

Animal

Body

Parts,

							Motor Vehicles; POS Tags and POS Trigrams,Pre-trained 300-dimensional GloVe embeddings	
7	clickbait	Clickbait detection on WeChat: A deep model integrating semantic and syntactic information	Liu et al., 2022	Proposed an MFWCD deep learning model that detects clickbait titles on the WeChat platform, constructs the first Chinese clickbait dataset and verifies the validity and interpretability of the model.	WeChat Clickbait Dataset	MFWCD (Multiple Features for WeChat Clickbait Detection), MFWCD-BERT, MFWCD-BiLSTM, K-Nearest Neighbor (KNN),	Semantic Features: Extracted using BERT (Bidirectional Encoder Representations from Transformers), Bi-LSTM	Accuracy F1 Precision Recall

						Random Forest (RF), Bernoulli Naive Bayes (NB), Support Vector Machines (SVM), Logistic Regression (LR), Bi-LSTM-A, Bi-GRU-A, Text-CNN,the base BERT model	(Bidirectional Long Short-Term Memory); Syntactic Features: Graph Attention Network (GAT); Part-of-speech tags and dependency; Auxiliary Features: metadata	
8	clickbait	Clickbait Detection with Style-aware	Wuy et al., 2020	The proposed a proven effectiveness SATC model that combines the	Webis Clickbait Corpus 2017	SATC (Style-aware Title Modeling	Content Features by Transformer; Title	Accuracy, Precision, Recall,

		Title Modeling and Co-attention		semantic and stylistic features of the title to improve the accuracy.	dataset, Fake News Challenge (FNC) Dataset^2	and Co-attention)	Stylistic Patterns by character-level Transformer (such as capitalisation, numeric characters, and punctuation); Title-Body Matching	F1 Score; Ablation Studies; Comparison with Baseline Models: CNN, LSTM, GRU-Att, and SiameseNet
9	clickbait	CLICK-ID: A novel dataset	William and Sari, 2020	Build a dataset of Indonesian clickbait titles to fill the gap in	CLICK-ID Dataset	Human annotation, Inter-Annotator	Headline Attributes: Original Headline,	Fleiss' Kappa score,

		for Indonesian clickbait headlines		Indonesian natural language processing.		Agreement, Bi-LSTM (Bidirectional Long Short-Term Memory), CNN (Convolutional Neural Network)	Publisher Information, Publication Date and Time, Category and Sub-Category, Clickbait Label	Avg Acc
10	clickbait	Does Clickbait Actually Atract More Clicks? Three Clickbait	Molina et al., 2021	https://blog.chartbeat.com/2015/11/20/youll-never-guess-how-chartbeats-data-scientists-came-up-with-the-single-greatest-headline/	Clickbait Headlines Dataset	2 deep learning models, Naïve Bayes, Support Vector Machine (SVM)	Linguistic Features: Questions, Lists, Wh Words, Demonstrative Adjectives, Positive Superlatives, Negative	Classification Agreement, Engagement Metrics, Negative Binomial Regres

							Superlatives, Modals	sions, Tukey HSD for Post-Hoc Comparisons
11	clickbait	From Clickbait to Fake News Detection_An Approach based on Detecting the Stance of Headlines to Articles	Bourgonje, Schneider and Rehm, 2017	Aim to detect clickbait and fake news by identifying positional consistency between news headlines and article content.	Fake News Challenge (FNC-1)	Logistic Regression Classifier, Binary Classifiers, Mallet’s Logistic Regression	n-grams (where n = 1..6), lemmatization, the length and inverse document frequency (IDF) of the n-grams, presence of question marks, negation	Relatedness score, Three-class score, Weighted score

							dependencies, Dependency parsing, sentence structures, and semantic similarity measures between headlines and article bodies	
13	clickbait	Investigating clickbait in Chinese social media: A study of WeChat	Zhang and Clough, 2020	Provide a clickbait detection method for Chinese social media by feature engineering and machine learning models.	WeChat Articles Dataset	MLP, feedforward Probabilistic Neural Network (PNN), Logistic Regression, Naïve Bayes,	Punctuation Usage, Word Usage, Clickbait Indicators, Metadata, SimHash	Cohen's Kappa, Precision, Recall, Accuracy, F1-measure,

						Random Forest, Support Vector Machine (SVM), Gradient Boosted Decision Tree (GBDT)		ROC-AUC.
14	sensation	Automatically quantifying the scientific quality and sensationalism of news records mentioning	Hoffmana and Justicz, 2016	Develop an automatically quantifying scientific quality and sensationalism in news reports, providing the possibility for automated text analysis and health news report assessment in the future.	10,700 records retrieved from Lexis Nexis database	maximum entropy model	Document-Term Matrix, Relevance Correlations, Scientific Quality Correlations, Sensationalism Correlatio	Accuracy, Cohen's Kappa, Intraclass Correlation Coefficient (ICC), Fleiss'

		pandemics_validating a maximum entropy machine-learning model					ns (Exposing, Speculating, Generalizing, Warning, and Extolling by five-point Likert scale)	Kappa, Krippendorff's Alpha, t-test
15	sensation	Indonesian Muslim society's reception of sensation language and invitation to	Astari et al., 2023	Discuss the reaction of the use of sensational language on the social media among the Indonesian Muslim societies to promote polygamy, suggesting that the sensational	14 respondents filled out the questionnaire	Qualitative Analysis	Language Features: Sensational phrases, Reception Typologies: Dominant Hegemonic Reception,	Public Reception Mapping

		polygamy on social media		language can increase the rejection of polygamy in some situations.			Negotiated Reception: Conditional acceptance, Opposition Reception: Rejection	
16	sensation	Mass media, the ‘sensational message’, and metamorphic truths	Vista, 2014	Explore sensationalism in mass media, especially how virus propagation phenomena, explaining the impact and changes in public recognition in the digital environment.	Susan Boyle Phenomenon video, Genovese Incident news report	Network Dynamics, Qualitative Analysis	Sensational Language	Viral Spread (such as the number of views or the rapidity with which a message spread

								s across networks)
17	sensation	Sensationalism in television news: A Review	Kleemans and Vettehen, 2009	Discuss sensationalism intelevision news and analyze the development reasons and the impact to the audiences, summarizing how sensational news attract veiewers' attention and affects their information processing.	2 online news articles, High Arousal and Low Arousal, each containing approximately 800-900 words	analyze television news content of the United States and Europe over different periods through various studies imited Capacity Model of Mediated Message Processing	Sensational Story Topics, Embedded Sensationalism, Sensational Pictures, Individual Case Histories, Number of Camera Shots, Decorative Editing Techniques, Music and Sound Effects,	Attention and Arousal, Memory and Recall, Audience Appeal,

							Eyewitness Camera, Close-Ups of Human Faces,	
18	sensation	The Effects of Sensational Language in News on Memory and Attitudes	Bell, 2015	Explore the impact of sensational language in the news report on the reader's attitude and memory, finding the changes in that high arousal language could affect viewers' attitude to the news content and exploring how the language selection forms audiences' recognition and emotional reactions.	127 students read designed high-arousal (sensational) and low-arousal (calm) versions of two New York Times	Counterbalanced, double-blind experimental design	40 Manipulated Terms, Attitude Measures, Knowledge Measures: Recognition Memory and Story Comprehension	Attitude Change Scores, Knowledge Scores, Manipulation Checks

					articles.			
19	sensation	Why humans value sensational news: An evolutionary perspective	Davis and McLeod, 2003	Explores why people continue to pay attention to sensational news, finding certain topics could be related to the core issues in human revolution psychology and illustrate that sensational news may satisfy the information needed in human evolution.	Newspaper Stories of 736 sensational front-page newspaper stories collected from eight countries (Australia, Bangladesh,	Human Coding, Kendall's Coefficient of Concordance, Category Frequency and Ranking, Stability Over Time	Story Categories: Murder/physical assault, Robbery/vandalism, Accidental/natural injury/death, Altruism/heroism, Suicide/self-inflicted injury, Abandoned/destitute family, Harm to a	Category Frequency and Ranking, Stability Over Time

					Canada, England, France, Germany, Mauritius, and the United States) spanning from 1700 to 2001.		child, Sexual assault/rape, Taking a stand/fighting back, Reputation, Marital/courtship anomalies, Miscellaneous stories	
20	sensation	Are `Sensational' News Stories	Uribe and Gunter, 2007	Explore the differences in emotion-aroused between sensational news	80 weekday newscasts,	Human Coding, Shot-Level Analysis	Emotion-Eliciting Content: Sex, Violence,	Overall Emotionality,

		More Likely to Trigger Viewers' Emotions than Non-Sensational News Stories: A Content Analysis of British TV News		and non-sensational news, finding that sensational news tends to contain visual and text content that is more likely to include emotion-aroused elements.	comprising 40 editions each from ITV and BBC1.		Destruction, Humour, Celebrities, Other Emotional Content	Visual and Verbal Emotionality, Chi-square tests
21	sensation	Proving the Obvious? What Sensationalism Contributes to the Time	Vettehen and Kleemans, 2017	Investigate how sensational news affects the length of audiences' view time, discovering that negative content and tabloid-style news packaging	190 participants watch up to 16 news stories.	Human Coding and Content Analysis	Negative Content: such as crime, accidents, or disasters, Tabloid Packaging,	Viewing Time, Analysis of Variance (ANOVA)

		Spent on News Video		significantly increased viewing time.			Viewing Time	
22	sensation	Sensationalism in news coverage: A comparative study in 14 television systems	Arbaoui, De Swert and van der Brug, 2020	A comparison of the sensationalization of television news in a cross-national manner, discovering that television systems that rely on business income and highly dispersed audiences are more likely to report sensational news.	812 broadcasts and 13,444 news items from 29 daily newscasts on public and private television stations across 14	Human Coding, intercoder reliability tests (Krippendorff's alpha and Cohen's kappa)	Sensational News Topics: crime, corruption, misconduct, violence, disasters, accidents, terrorism, sex, drugs, celebrities; Storytelling Sensationalism: "ordinary actors";	Sensationalism Categories: sensational news topics, sensational storytelling, and sensational formal features; Intercoder

					television systems.			Reliability, Multilevel Logistic Regression
23	sensation	Towards a pragma-linguistic framework for the study of sensationalism in news headlines	Molek-Kozakowska, 2013	Aims to build a pragma-linguistic framework, reveals how common linguistic strategies in news headlines could enhance sensational effect and helps identify sensational language features in news reports.	120 entries (headlines, subheadlines, and lead-ins) of the British newspaper Daily Mail	Human Coding and Focus Groups, Pragma-linguistic Analysis	Illocutions, Semantic Macrostructures/Themes, Narrative Formulas, Interpersonal and Textual Devices	Sensationalism Ratings Five-point Likert scale, Focus Group Discussions

2.2. Gaps in the Literature

Current sensational language research is more likely to be conducted in a corpus analysis method, barely in a natural language processing method by programming. We expect to build the bridge by programming technique.

2.3. Relation to Current Study

The use of sensational language and clickbait in news headlines both aim to capture the audience's attention. We investigate different training models and features of clickbait in different languages, hoping to get some ideas that could inspire us in sensational language research.

Chapter 3: Methodology

3.1. Dataset Description

MIRUKU sensational news headline dataset.

We create a MIRUKU sensational news headline dataset. The dataset is made based on the News Clickbait Dataset from Kaggle (Singh, 1996; Chakraborty et al., 2016).The clickbait corpus consists of article headlines from 'BuzzFeed', 'Upworthy', 'ViralNova', 'Thatscoop', 'Scoopwhoop' and 'ViralStories'. The non-clickbait article headlines are collected from 'WikiNews', 'New York Times', 'The Guardian', and 'The Hindu' (Singh, 1996).

MIRUKU sensational news headline dataset contains 30,425 rows and 10 columns: headline, clickbait, sensation, sensation_score, sensation_reason, emotion, arousal, arousal_score, arousal_reason, arousal_category. The MIRUKU dataset is annotated with the help of OpenAI GPT-4o. We adopt the Likert scale as guidance. The guidance is originally for human annotators to score the sensation in the news headline (Molek-Kozakowska, 2013).

The MIRUKU sensational news headline dataset is created in two steps:

1: Annotate for sensation, sensation_score, sensation_reason. We delete 117 rows with multiple sensation indicators (multiple 1 and 0). The dataset remains 31,883 rows.

2: Annotate for emotion, arousal, arousal_score, arousal_reason, arousal_category. We delete 1,459 rows with multiple sensation indicators (multiple 1 and 0). The dataset remains 30,424 rows.

Table 2

headline	clickbait	sensation	sensation_score	sensation_reason	emotion	arousal	arousal_score	arousal_reason	arousal_category
$500k Of Student Loan Debt In 37 Seconds	1	1	3.75	The use of a large monetary figure in conjunction with a very short timeframe is inherently dramatic and attention-grabbing. The	anger, sadness, fear	Yes	0.8	The mention of substantial debt accumulated in a very short period creates a sense of urgency and distress, which contributes to	Fairly

	text aims to provoke shock and curiosity .	high arousal levels.

Superlative adjective words list (Biyani, Tsioutsiouliklis and Blackmer, 2016).

We create a superlative adjective word (594 words) list in five steps for experiment purposes. The sources of the list come from:

1. Internet.

2. Extract superlative adjectives and adverbs from NLTK's built-in corpora, including Brown, webtext, Reuters, Movie_Reviews, and Gutenberg.

We utilize Regular Expression to define the regular superlative patterns. For short adjectives and adverbs, we applied "-est" and "iest" suffixes. We applied most, least, best, and worst for long adjectives and adverbs before the adjective. For common and irregular superlatives adjectives and adverbs, we add best, worst, furthest, farthest', least, most, latest, last, nearest, and dearest.

3. Hallman (2016) considers that 'all' and 'every' as quantity superlative formation and add them to the list at the final step.

4. To reduce distractions, we decide to remove the superlative adverb list.

5. Manual inspect each superlative adjective word. 594 words were eventually left in the list.

3.2. Model and Algorithm Development

We first extract 11 features and train each selected feature separately, and then we concatenate all features. Before we trained all features, we used the Principal Component Analysis (PCA) on different emotions in each news headline to determine the threshold with sensation score and arousal score. We processed the feature reduction in 2 steps. First, we applied the filter method, the variance threshold, to eliminate constant or nearly constant features to reduce dimension. By doing so, we can reduce the risk of overfitting, save computing resources, and improve model interpretability. The threshold we chose is 0.001 and 1,236 features remained. Second, we applied the wrapper method, Cuckoo Search via Lévy flights (Sangwan and Bhatia, 2020). The advantage of Cuckoo Search via Lévy flights is that we can set the objective function as f1-score, which is the evaluation metric we mainly focus on. We can find global optimal instead of local optimal and handle non-linear co-relationship features effectively. D-BullyRumbler a safety rumble strip to resolve online denigration bullying using a hybrid flter-wrapper approach, Cuckoo Search via L´evy Flights. Furthermore, we chose XGBoost as our baseline model and AdaBoost, CATBoost and Random Forest as our experiment model. We select Random Search because of the limitation of computing power. Compared with Grid Search, Random Search is more efficient in high-dimension hyper parameter search space because Random Search can cover more dimensions on average, without wasting resources on unimportant dimensions, thereby performing higher computing efficiency. (Bergstra and Bengio, 2012). We use K-Fold cross-validation (n_splits=5) and evaluation metric F1 score because we want our model to treat sensational or non-sensational equally. Lastly, we analyse the feature importance and the SHAP ((SHapley Additive exPlanations) dependency, establishing the fundamentals for future research.

Chapter 4: Implementation and Results

4.1. Model Implementation

Experiments are compeleted on Google Colab Pro+, with A100 GPU, System RAM 83.5 GB, GPU RAM 40.0 GB, Disk 201.2 GB.

Parameter Type	XGBoost	AdaBoost	CatBoost	Random Forest
Number of estimators	'n_estimators': [100, 300, 500]	n_estimators': [100, 300, 500]	iterations': [100, 300, 500]	n_estimators': [100, 300, 500]
Learning rate	'learning_rate': [0.01, 0.1, 0.3]	learning_rate': [0.01, 0.1, 0.3]	learning_rate': [0.01, 0.1, 0.3]	N/A
Tree depth	'max_depth': [3, 6, 9]	base_estimator__max_depth': [3, 6, 9]	depth': [3, 6, 9]	max_depth': [3, 6, 9]
Minimum samples per split	'min_child_weight': [1, 3]	base_estimator__min_weight_fraction_leaf': [0, 0.1]	l2_leaf_reg': [1, 3]	min_samples_leaf': [1, 3]
Feature subsampling	'colsample_bytree': [0.8, 1.0]	base_estimator__max_features': [0.8, 1.0]	colsample_bylevel': [0.8, 1.0]	max_features': [0.8, 1.0]
Row subsampling	'subsample': [0.8, 1.0]	N/A	subsample': [0.8, 1.0]	bootstrap': [True, False]
Regularization	'gamma': [0, 0.1]	N/A	random_strength': [0, 0.1]	min_impurity_decrease': [0, 0.1]
Model-specific	N/A	algorithm': ['SAMME', 'SAMME.R']	N/A	N/A

Stage 1: Train with each feature separately: 1.Number of words, 2.Number of stop words, 3.Ratio of stop words to content words, 4.Flesch-Kincaid Readability, 5.Subjectivity and Objectivity, 6.Sentiment analysis (Negative Sentiment, Neutral Sentiment, Positive Sentiment, Compound Sentiment), 7.Elongated Words, 8.Punctuation (Currency symbols, Exclamation marks, Question marks, Ellipsis, Emphasis marks, Multiple exclamation marks, Single quotes, Double quotes, Contracted word forms), 9.TF-IDF with Stop words, 10.TF-IDF without Stop words

11.Syntactic 4-grams

Stage 2:

Calculate the Optimal Threshold for baseline model XGBoost. First use Principal component analysis (PCA) on emotion column and then combine the sensational score and arousal score by using XGBClassifier.

Concatenate all features (except feature TF-IDF without Stop words) and train each model separately: XGBoost, XGBoost with Superlative Adjective Words List, XGBoost with Threshold, XGBoost with Superlative adjective words list and threshold, AdaBoost, CATBoost, Random Forest, CATBoost on Test set.

Stage 3: SHAP analysis.

4.2. Results

Train with each feature separately:

1.Number of words (Chakraborty et al., 2016), (Indurthi et al., 2018)

2.Number of stop words (Biyani, Tsioutsiouliklis and Blackmer, 2016; Chakraborty et al., 2016; Indurthi et al., 2018)

3.Ratio of stop words to content words (Chakraborty et al., 2016)

4.Flesch-Kincaid Readability (Biyani, Tsioutsiouliklis and Blackmer, 2016)

5.Subjectivity and Objectivity (Biyani, Tsioutsiouliklis and Blackmer, 2016; Molek-Kozakowska, 2013; Volkova et al., 2017).

6.Sentiment analysis (Biyani, Tsioutsiouliklis and Blackmer, 2016; Molek-Kozakowska, 2013; Brown et al., 2018; Yang, Mukherjee and Gragut, 2017; Chakraborty et al., 2016), including Negative Sentiment (Arlim et al., 2022; Cano Mora, 2009; Oraby et al., 2016; Buschmeier, Cimiano and Klinger, 2014; Uribe and Gunter, 2007; Vettehen and Peeters, 2008; Vettehen and Kleemans, 2017; Molek-Kozakowska, 2013; Brown et al., 2018); Neutral Sentiment; Positive Sentiment; Compound Sentiment.

7.Elongated Words (Biyani, Tsioutsiouliklis and Blackmer, 2016; Arlim et al., 2022)

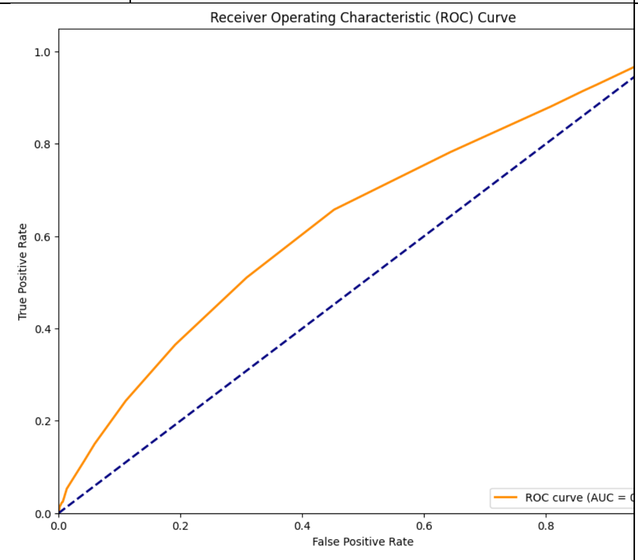
8.Punctuation (Biyani, Tsioutsiouliklis and Blackmer, 2016; Indurthi et al., 2018; Zhang and Clough, 2020; Chakraborty et al., 2016; Arlim et al., 2022), including Currency symbols; Exclamation marks (Biyani, Tsioutsiouliklis and Blackmer, 2016; Arlim et al., 2022; Karoui et al., 2016); Question marks (Biyani, Tsioutsiouliklis and Blackmer, 2016; Arlim et al., 2022); Ellipsis (Chakraborty et al., 2016); Emphasis marks **(**Chakraborty et al., 2016; Bharti et al., 2016); Multiple exclamation marks (Chakraborty et al., 2016; Bharti et al., 2016; Karoui et al., 2016); Single quotes (Indurthi et al., 2018; Buschmeier, Cimiano and Klinger, 2014); Double quotes (Indurthi et al., 2018; Bharti et al., 2016; Buschmeier, Cimiano and Klinger, 2014); Contracted word forms (Biyani, Tsioutsiouliklis and Blackmer, 2016); 9.TF-IDF with Stop words (Biyani, Tsioutsiouliklis and Blackmer, 2016); 10.TF-IDF without Stop words (Biyani, Tsioutsiouliklis and Blackmer, 2016; Indurthi et al., 2018); 11.Syntactic 4-grams (Chakraborty et al., 2016; Yang, Mukherjee and Gragut, 2017).

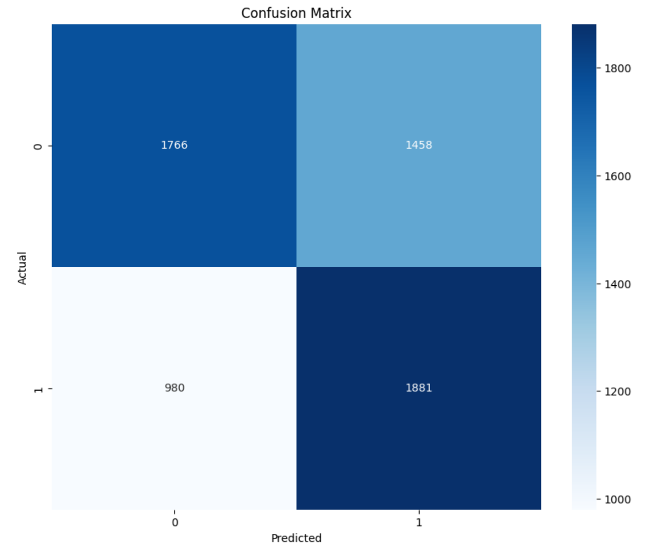
1.Number of words analysis:

Table 3

Numbe r of words in the	Best parameters: {'subsample': 0.8, 'n_estimators': 100, 'min_child_weight': 3, 'max_depth': 9, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}
	Best cross-validation score: 0.606811546536494
	Validation Set Classification Report:

headline		precision	recall	f1-score	support
	non-sensation	0.64	0.55	0.59	3224
	sensation	0.56	0.66	0.61	2861
	accuracy			0.60	6085
	macro avg	0.60	0.60	0.60	6085
	weighted avg	0.61	0.60	0.60	6085



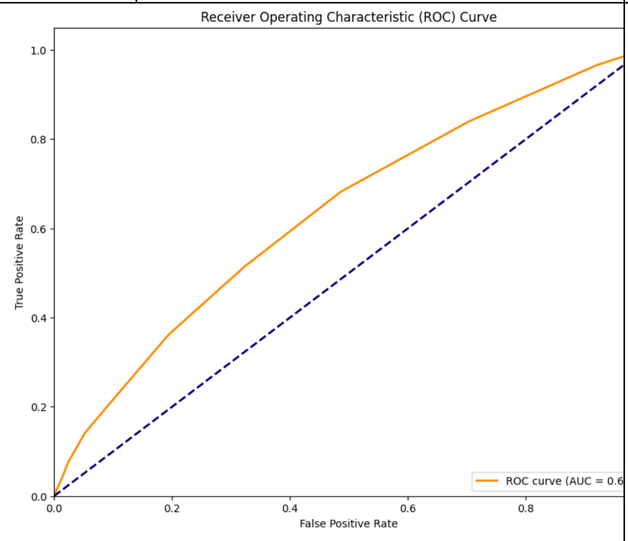


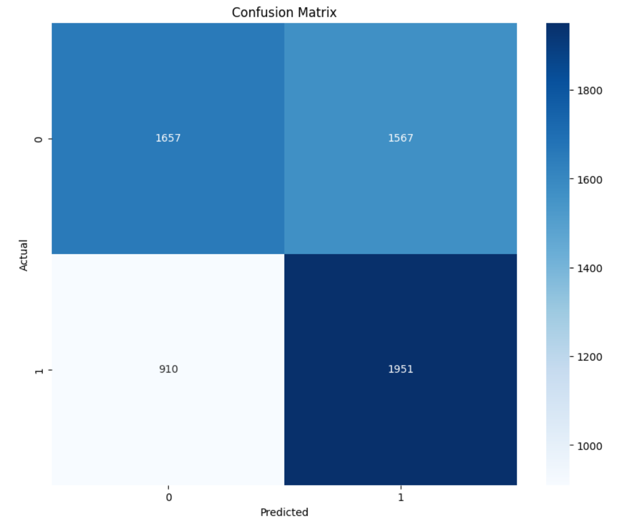
2.Number of stop words:

Table 4

Number of stop words in the headlines	Best parameters: {'subsample': 0.8, 'n_estimators': 100, 'min_child_weight': 3, 'max_depth': 9, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.6304409368430474				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.65	0.51	0.57	3224
	sensation	0.55	0.68	0.61	2861

	accuracy			0.59	6085
	macro avg	0.60	0.60	0.59	6085
	weighted avg	0.60	0.59	0.59	6085

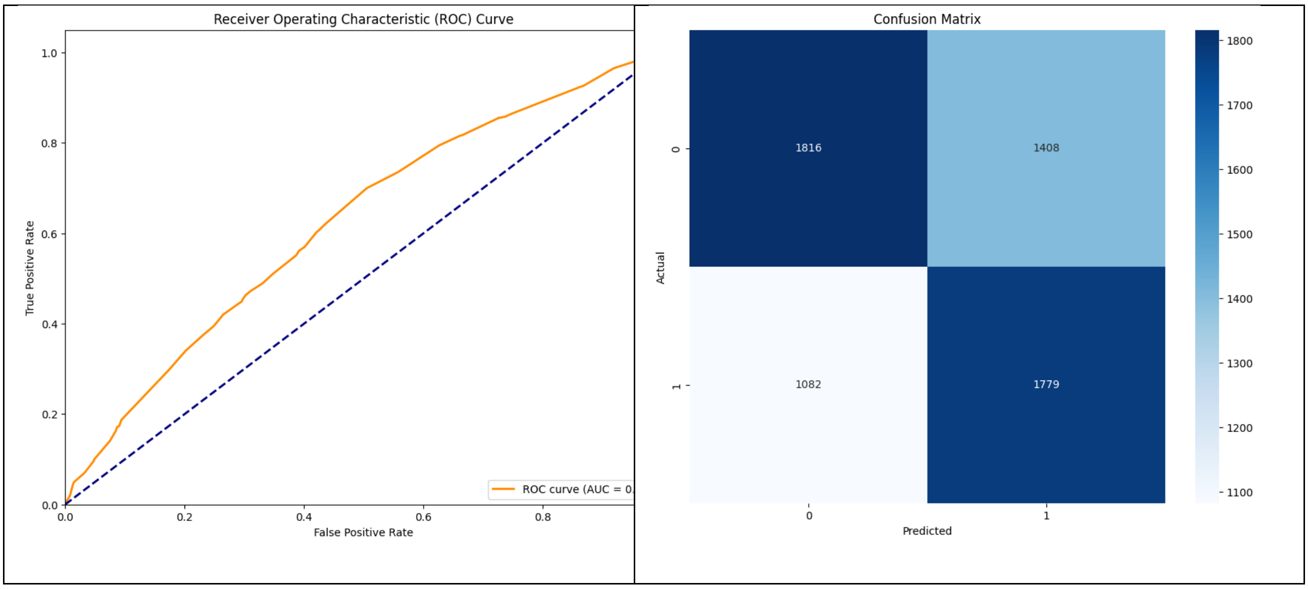




3.Ratio of stop words to content words:

Table 5

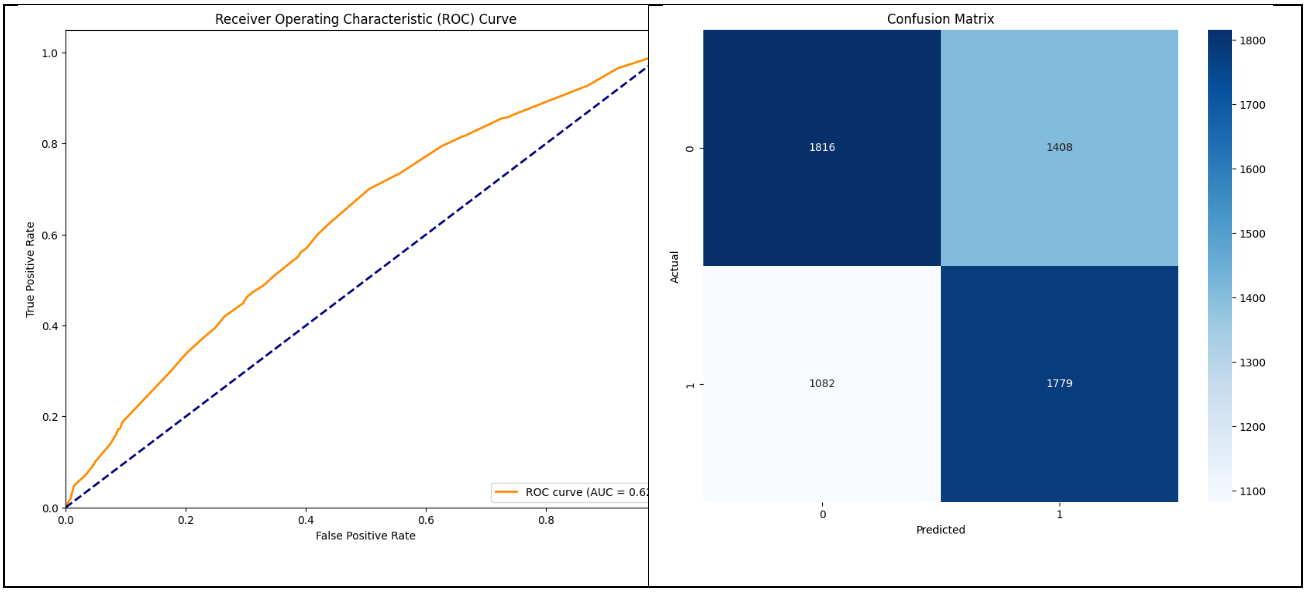
The ratio of the number of stop words to the number of content words	Best parameters: {'subsample': 1.0, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 3, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.608918170266611				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.63	0.56	0.59	3224
	sensation	0.56	0.62	0.59	2861
	accuracy			0.59	6085
	macro avg	0.59	0.59	0.59	6085
	weighted avg	0.59	0.59	0.59	6085



4.Flesch-Kincaid Readability

Table 6

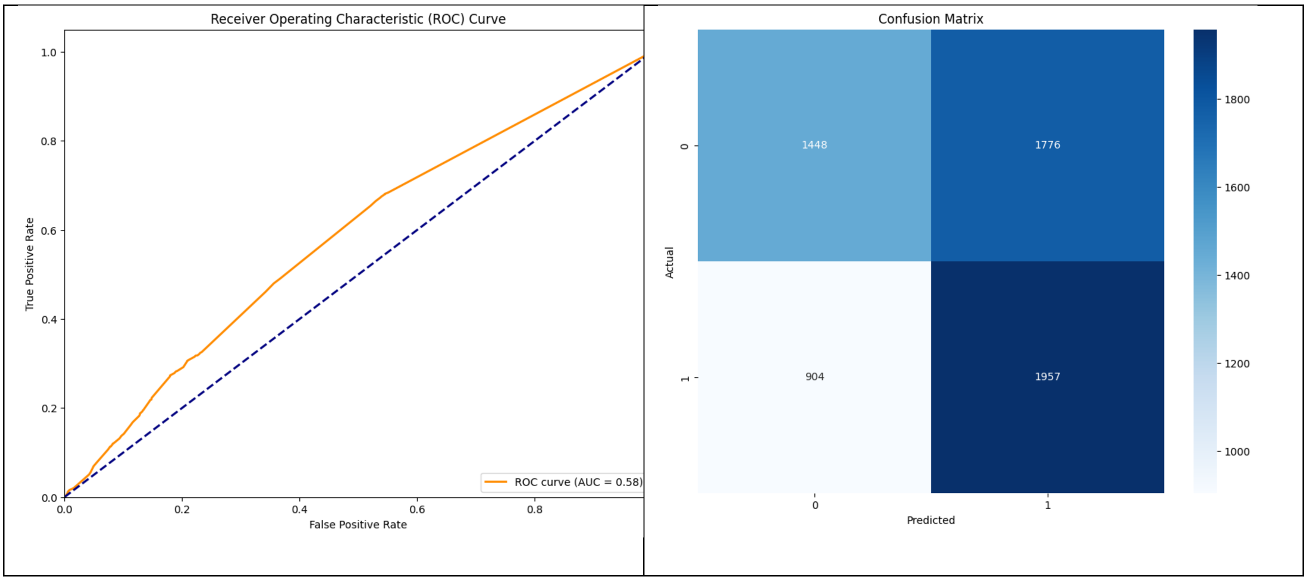
<table>
<tr><td rowspan="9">Informality (Flesch-Kincaid Readability)</td><td colspan="5">Best parameters: {'subsample': 1.0, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 3, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}</td></tr>
<tr><td colspan="5">Best cross-validation score: 0.608918170266611</td></tr>
<tr><td colspan="5">Validation Set Classification Report:</td></tr>
<tr><td></td><td>precision</td><td>recall</td><td>f1-score</td><td>support</td></tr>
<tr><td>non-sensation</td><td>0.63</td><td>0.56</td><td>0.59</td><td>3224</td></tr>
<tr><td>sensation</td><td>0.56</td><td>0.62</td><td>0.59</td><td>2861</td></tr>
<tr><td>accuracy</td><td></td><td></td><td>0.59</td><td>6085</td></tr>
<tr><td>macro avg</td><td>0.59</td><td>0.59</td><td>0.59</td><td>6085</td></tr>
<tr><td>weighted avg</td><td>0.59</td><td>0.59</td><td>0.59</td><td>6085</td></tr>
</table>



5.Subjectivity and Objectivity:

Table 7

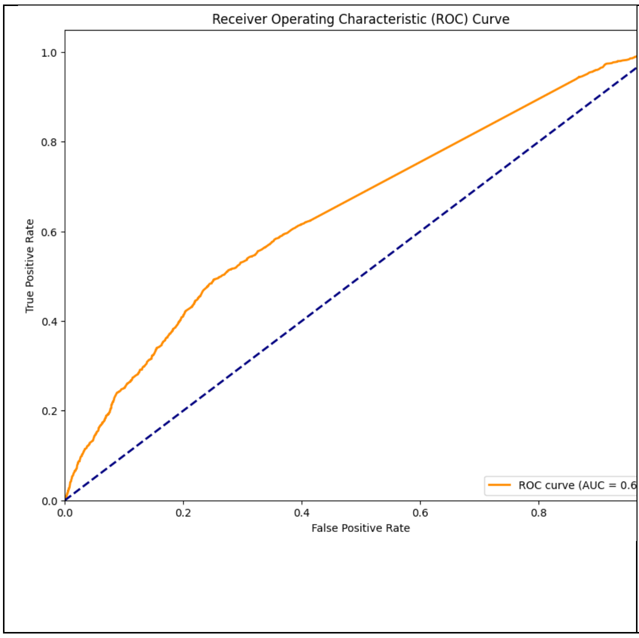
Sentence Subjectivity and Objectivity Evaluation	Best parameters: {'subsample': 0.8, 'n_estimators': 100, 'min_child_weight': 1, 'max_depth': 3, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.6146555909839979				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.62	0.45	0.52	3224
	sensation	0.52	0.68	0.59	2861
	accuracy			0.56	6085
	macro avg	0.57	0.57	0.56	6085
	weighted avg	0.57	0.56	0.55	6085

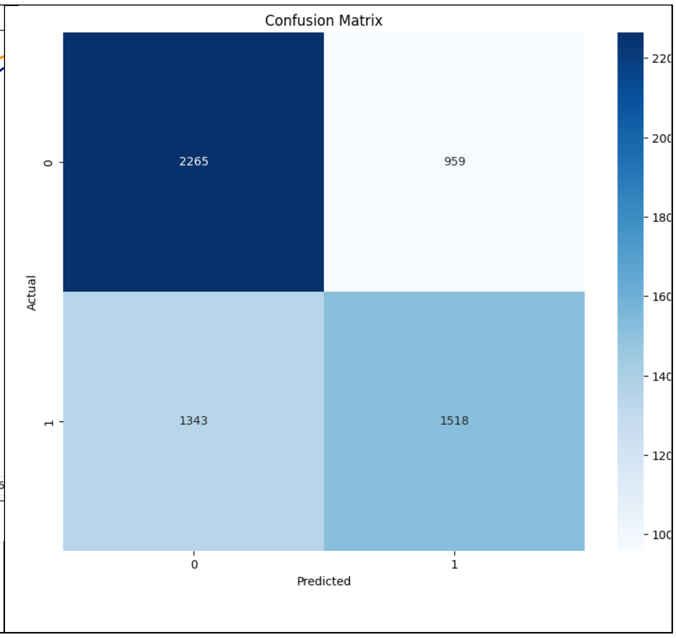


6.Sentiment analysis: Negative Sentiment, Neutral Sentiment, Positive Sentiment, Compound Sentiment

Table 8

Sentiment Analysis	Best parameters: {'subsample': 1.0, 'n_estimators': 100, 'min_child_weight': 3, 'max_depth': 3, 'learning_rate': 0.1, 'gamma': 0.1, 'colsample_bytree': 1.0}				
	Best cross-validation score: 0.5853623871049011				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.63	0.70	0.66	3224
	sensation	0.61	0.53	0.57	2861
	accuracy			0.62	6085
	macro avg	0.62	0.62	0.62	6085
	weighted avg	0.62	0.62	0.62	6085

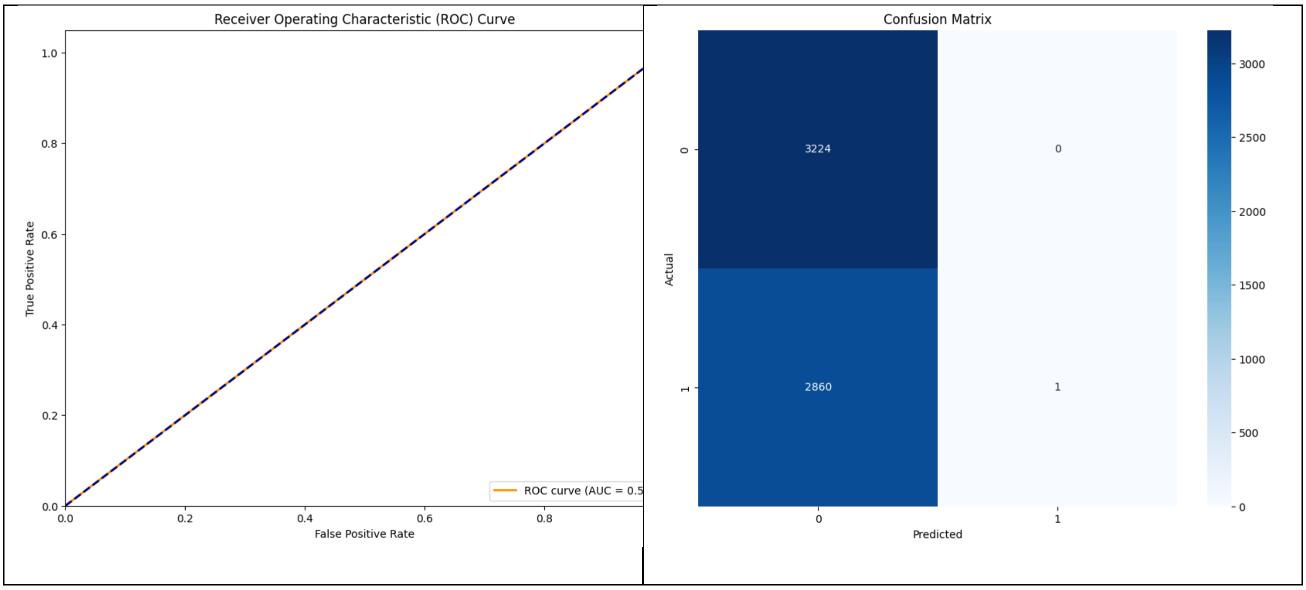




7.Elongated Words

Table 9

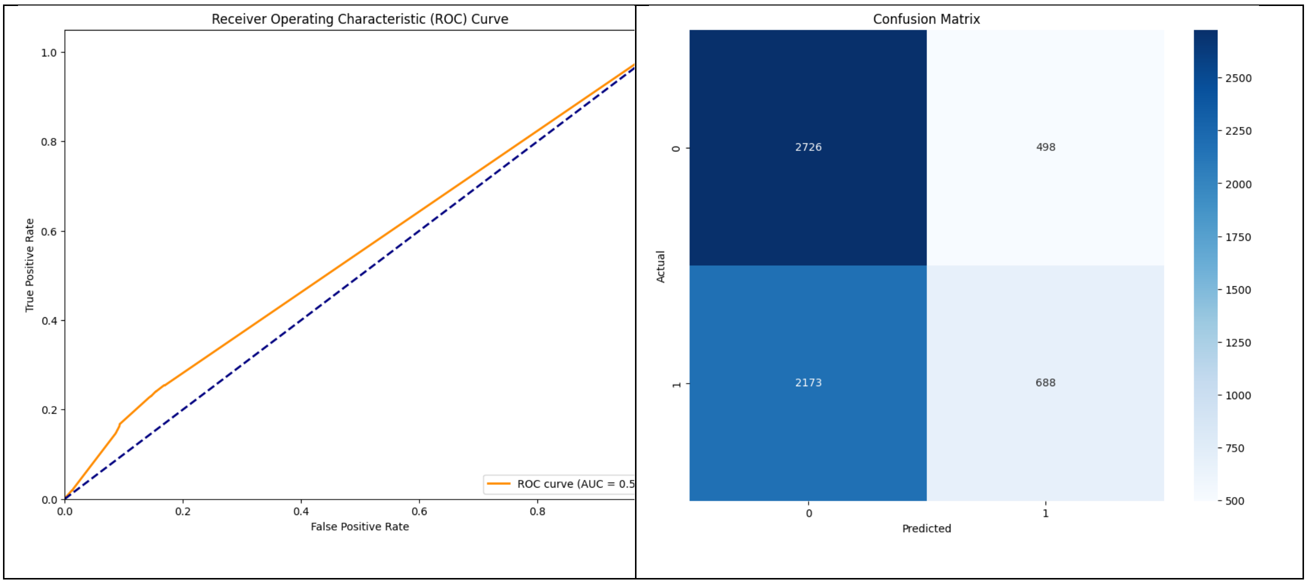
Elongated Words	Best parameters: {'subsample': 0.8, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 9, 'learning_rate': 0.3, 'gamma': 0.1, 'colsample_bytree': 1.0}				
	Best cross-validation score: 0.3975923623879674				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.53	1.00	0.69	3224
	sensation	1.00	0.00	0.00	2861
	accuracy			0.53	6085
	macro avg	0.76	0.50	0.35	6085
	weighted avg	0.75	0.53	0.37	6085



8.Punctuation: Currency symbols, Exclamation marks, Question marks, Ellipsis, Emphasis marks, Multiple exclamation marks, Single quotes, Double quotes, Contracted word forms.

Table 10

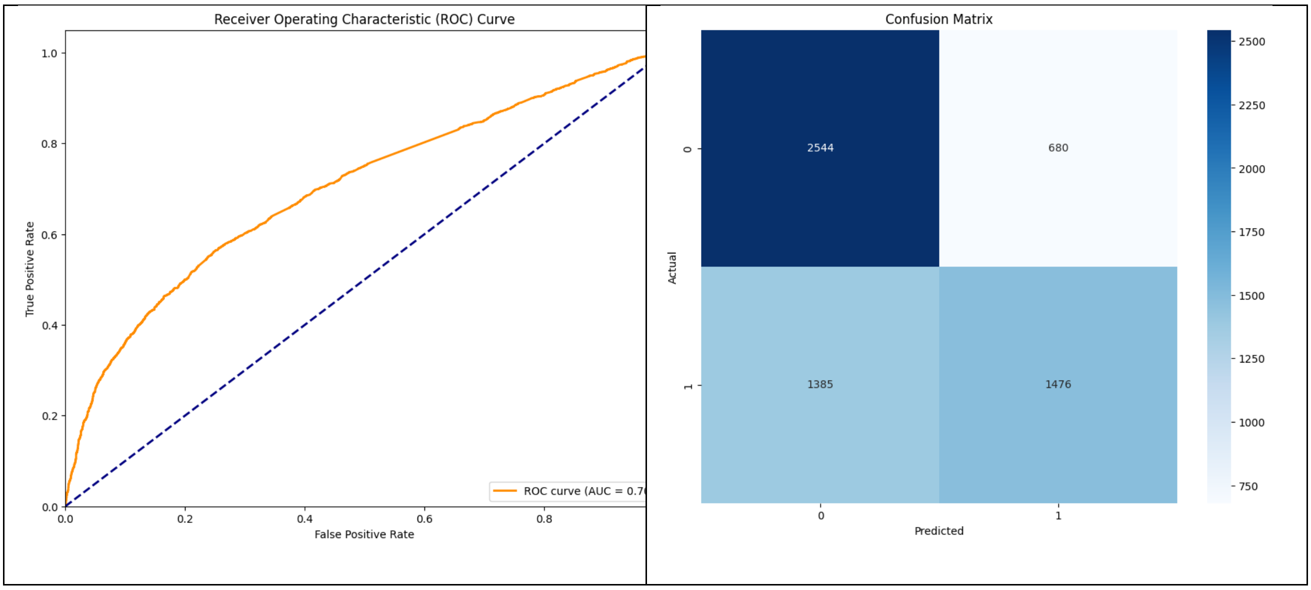
Punctuation	Best parameters: {'subsample': 1.0, 'n_estimators': 300, 'min_child_weight': 3, 'max_depth': 3, 'learning_rate': 0.1, 'gamma': 0.1, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.35970612589472656				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.56	1.00	0.67	3224
	sensation	0.58	0.00	0.34	2861
	accuracy			0.56	6085
	macro avg	0.57	0.54	0.51	6085
	weighted avg	0.57	0.56	0.52	6085



9.TF-IDF with Stop words

Table 11

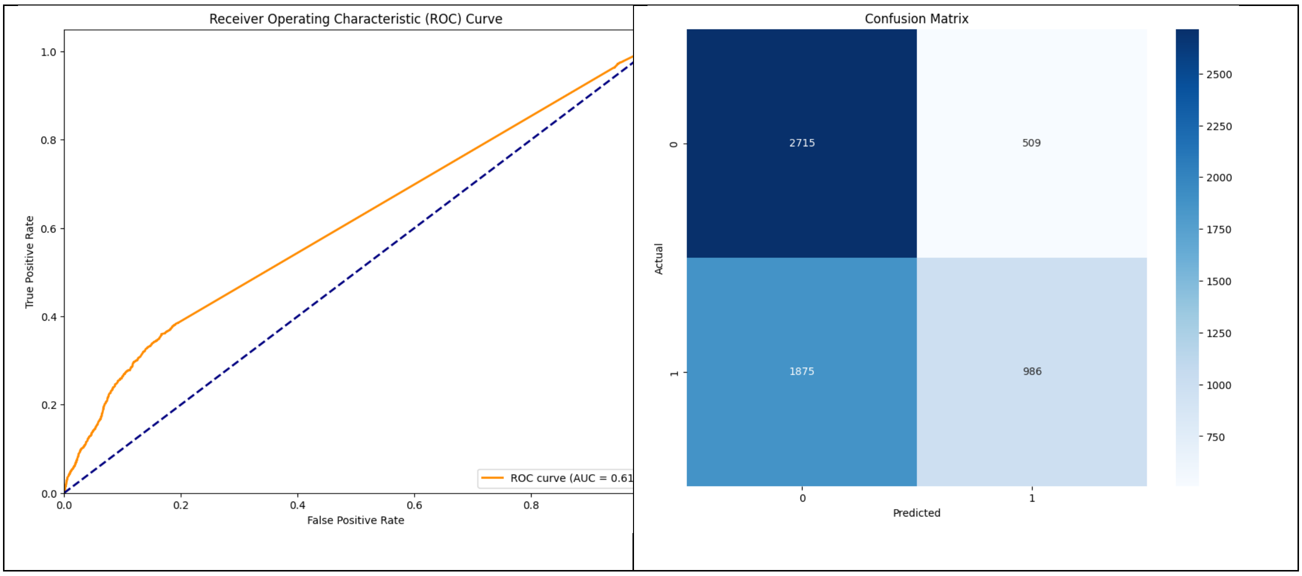
TF-IDF with stop words	Best parameters: {'subsample': 0.8, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 9, 'learning_rate': 0.1, 'gamma': 0.1, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.6168231099356203				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.65	0.79	0.71	3224
	sensation	0.68	0.52	0.59	2861
	accuracy			0.66	6085
	macro avg	0.67	0.65	0.65	6085
	weighted avg	0.68	0.66	0.65	6085



10.TF-IDF without Stop words

Table 12

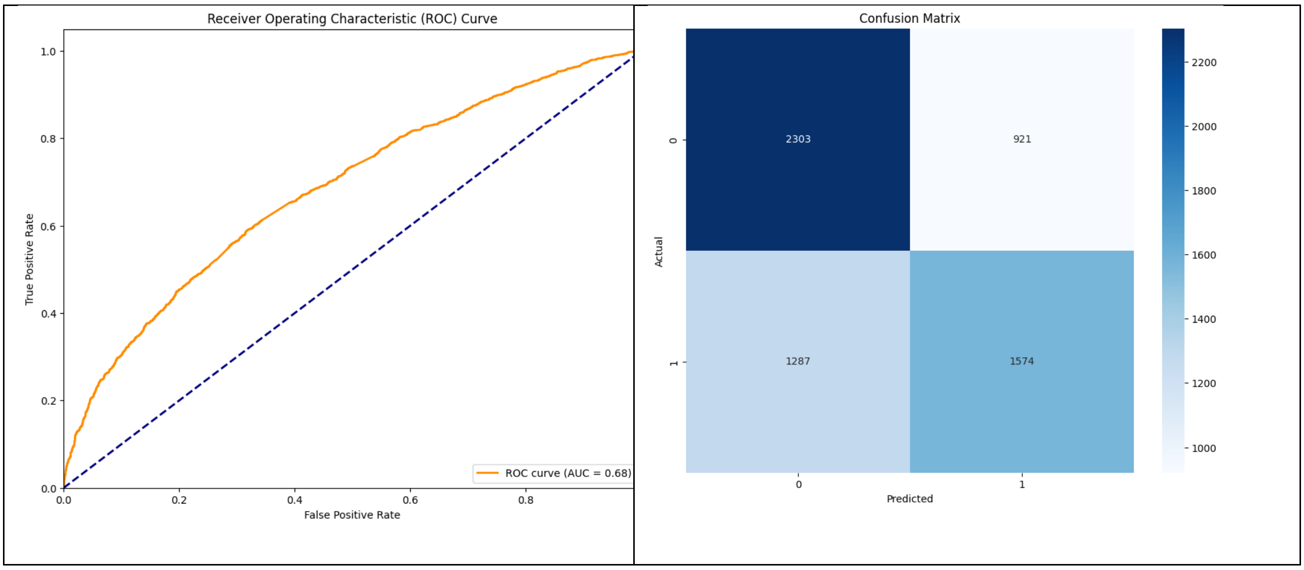
TF-IDF w/o stop words	Best parameters: {'subsample': 0.8, 'n_estimators': 300, 'min_child_weight': 3, 'max_depth': 3, 'learning_rate': 0.1, 'gamma': 0.1, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.5345758987416854				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.59	0.84	0.69	3224
	sensation	0.66	0.34	0.45	2861
	accuracy			0.61	6085
	macro avg	0.63	0.59	0.57	6085
	weighted avg	0.62	0.61	0.58	6085



11.Syntactic 4-grams

Table 13

Syntactic Ngrams	Best parameters: {'subsample': 0.8, 'n_estimators': 500, 'min_child_weight': 3, 'max_depth': 9, 'learning_rate': 0.1, 'gamma': 0, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.6232884322577538				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.64	0.71	0.68	3224
	sensation	0.63	0.55	0.59	2861
	accuracy			0.64	6085
	macro avg	0.64	0.63	0.63	6085
	weighted avg	0.64	0.64	0.63	6085



Stage 2:

Calculate the Optimal Threshold for baseline model XGBoost.

Concatenate all features (except feature TF-IDF without Stop words) and train each model separately: XGBoost, XGBoost with Superlative Adjective Words List, XGBoost with Threshold, XGBoost with Superlative adjective words list and threshold, AdaBoost, CATBoost, Random Forest, CATBoost on Test set.

Table 14

XGBoost with Superlative Adjective Words List	Best parameters: {'subsample': 0.8, 'n_estimators': 500, 'min_child_weight': 3, 'max_depth': 6, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.6644156657481959				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.68	0.67	0.68	3223
	sensation	0.64	0.65	0.64	2862
	accuracy			0.66	6085
	macro avg	0.66	0.66	0.66	6085
	weighted avg	0.66	0.66	0.66	6085

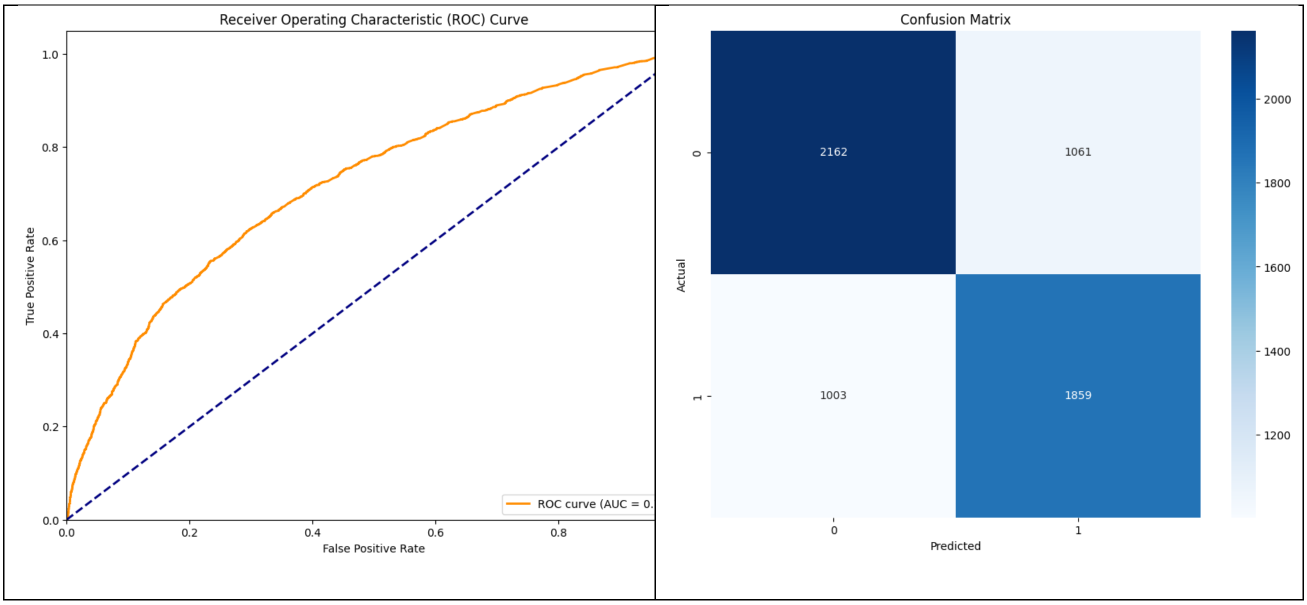


Table 15

XGBoost with Threshold	Best parameters: {'subsample': 0.8, 'n_estimators': 500, 'min_child_weight': 1, 'max_depth': 6, 'learning_rate': 0.01, 'gamma': 0.1, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.6634232268949644				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.70	0.59	0.64	3223
	sensation	0.61	0.72	0.66	2862
	accuracy			0.65	6085
	macro avg	0.65	0.65	0.65	6085
	weighted avg	0.66	0.65	0.65	6085

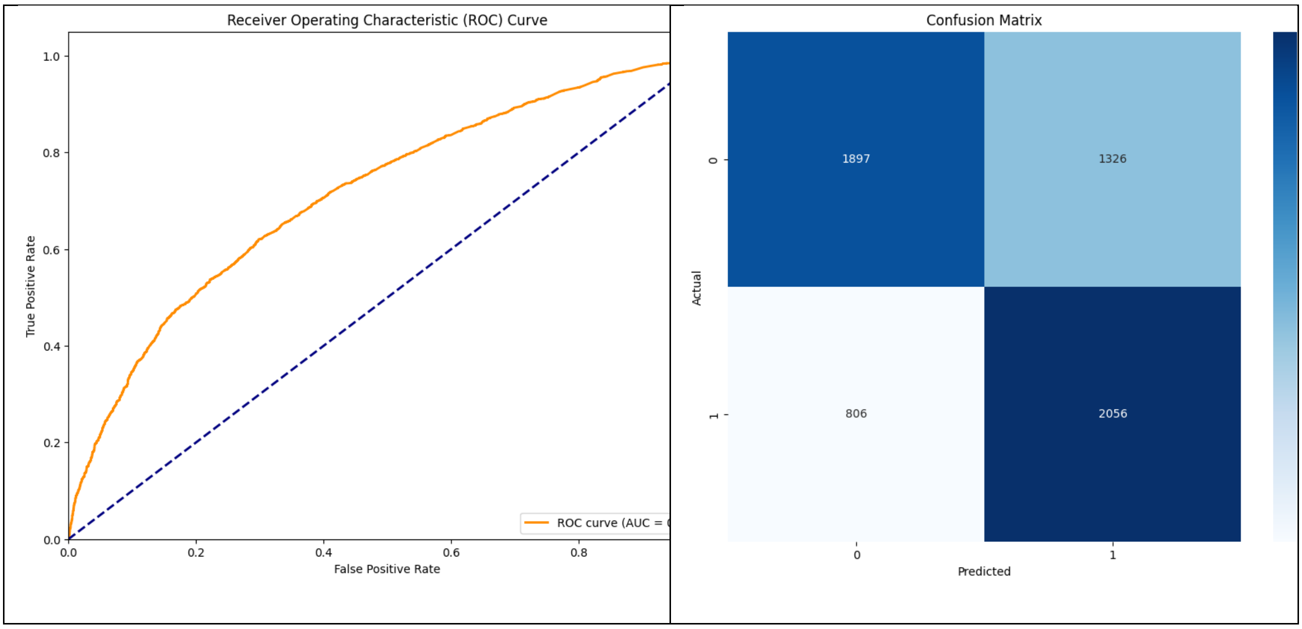


Table 16

XGBoost with Superlative adjective words list and threshold	Best parameters: {'subsample': 0.8, 'n_estimators': 500, 'min_child_weight': 3, 'max_depth': 6, 'learning_rate': 0.01, 'gamma': 0, 'colsample_bytree': 0.8}				
	Best cross-validation score: 0.6644156657481959				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.70	0.59	0.64	3223
	sensation	0.61	0.72	0.66	2862
	accuracy			0.65	6085
	macro avg	0.66	0.65	0.65	6085
	weighted avg	0.66	0.65	0.65	6085

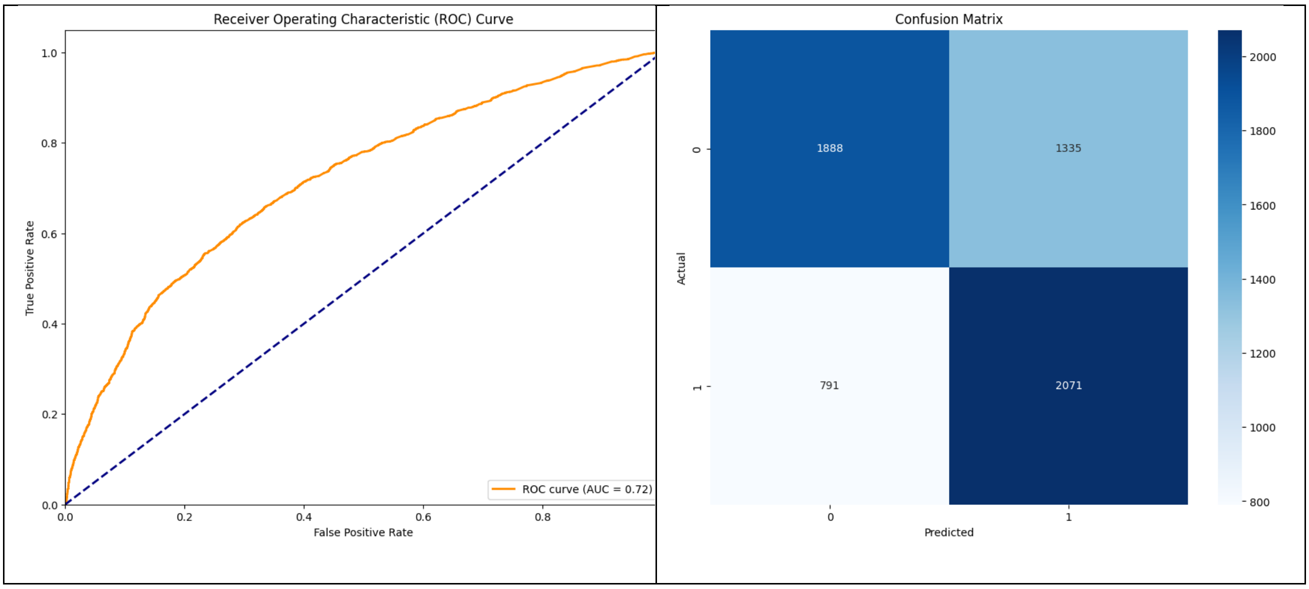


Table 17

<table>
<tr><td rowspan="9">AdaBoost</td><td colspan="5">Best parameters: {'n_estimators': 300, 'learning_rate': 0.01, 'base_estimator__min_weight_fraction_leaf': 0, 'base_estimator__max_features': 1.0, 'base_estimator__max_depth': 3, 'algorithm': 'SAMME.R'}</td></tr>
<tr><td colspan="5">Best cross-validation score: 0.6658196625136333</td></tr>
<tr><td colspan="5">Validation Set Classification Report:</td></tr>
<tr><td></td><td>precision</td><td>recall</td><td>f1-score</td><td>support</td></tr>
<tr><td>non-sensation</td><td>0.68</td><td>0.65</td><td>0.67</td><td>3223</td></tr>
<tr><td>sensation</td><td>0.62</td><td>0.66</td><td>0.64</td><td>2862</td></tr>
<tr><td>accuracy</td><td></td><td></td><td>0.65</td><td>6085</td></tr>
<tr><td>macro avg</td><td>0.65</td><td>0.65</td><td>0.65</td><td>6085</td></tr>
<tr><td>weighted avg</td><td>0.65</td><td>0.65</td><td>0.65</td><td>6085</td></tr>
</table>

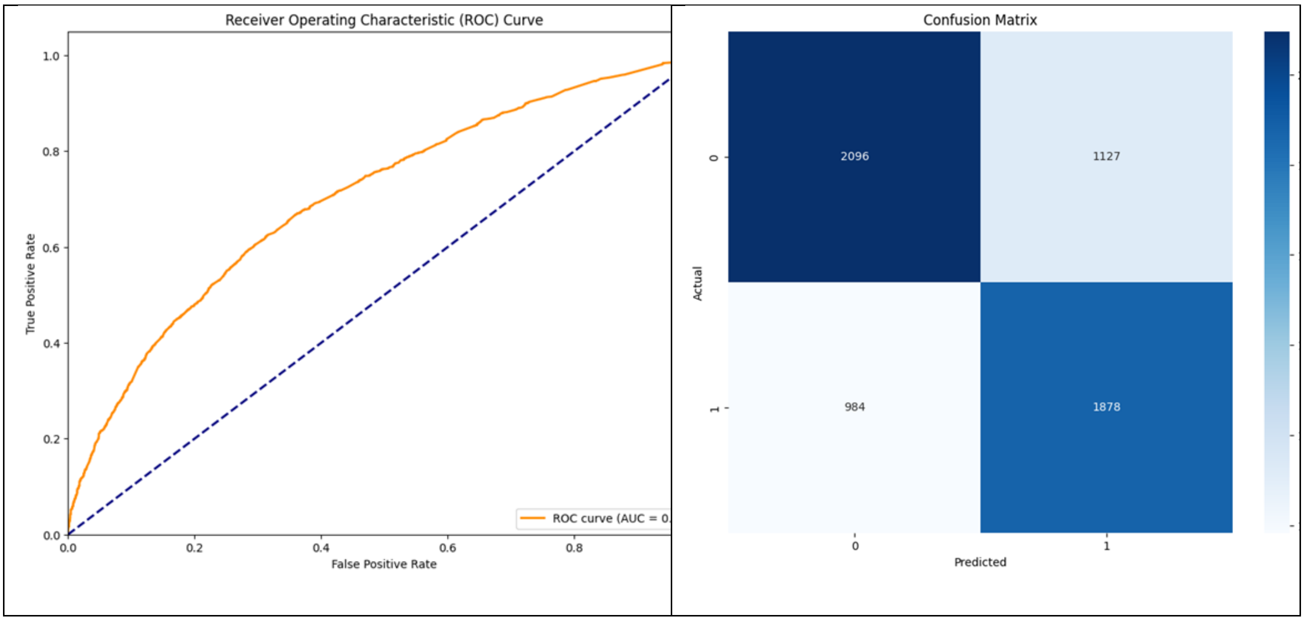


Table 18

CATBoost	Best parameters: {'subsample': 0.8, 'random_strength': 0.1, 'learning_rate': 0.1, 'l2_leaf_reg': 3, 'iterations': 100, 'depth': 9, 'colsample_bylevel': 1.0}				
	Best cross-validation score: 0.6697000335964166				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.68	0.68	0.68	3223
	sensation	0.64	0.64	0.64	2862
	accuracy			0.66	6085
	macro avg	0.66	0.66	0.66	6085
	weighted avg	0.66	0.66	0.66	6085

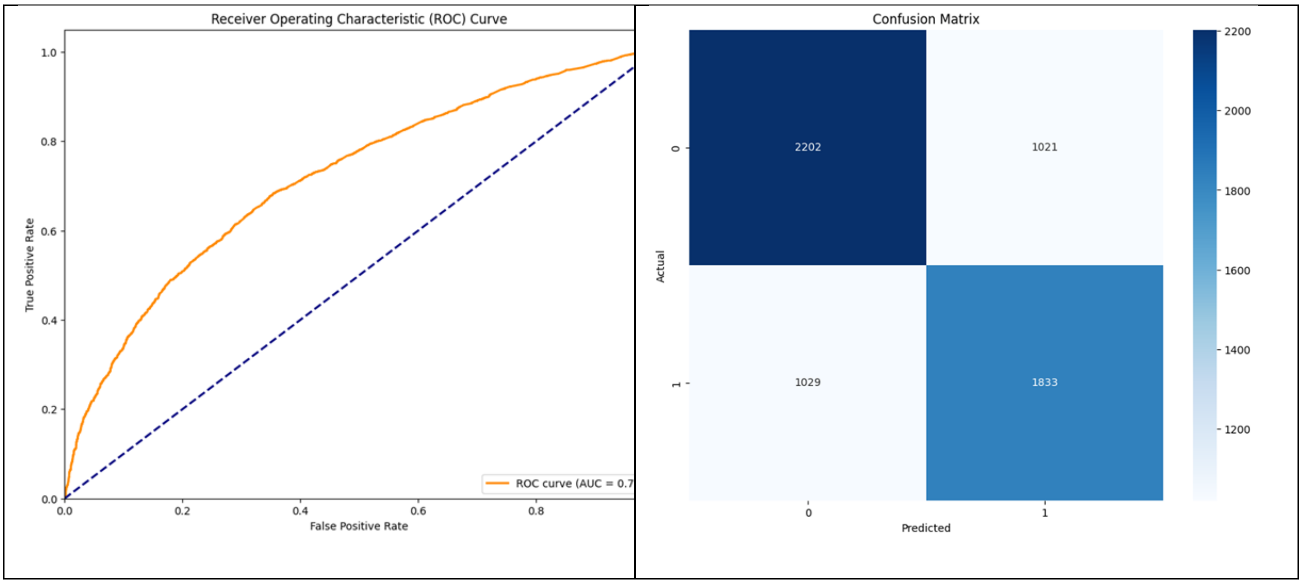


Table 19

Random Forest	Best parameters: {'n_estimators': 500, 'min_samples_leaf': 3, 'min_impurity_decrease': 0, 'max_features': 1.0, 'max_depth': 3, 'bootstrap': True}				
	Best cross-validation score: 0.6611619132764461				
	Validation Set Classification Report:				
		precision	recall	f1-score	support
	non-sensation	0.69	0.57	0.62	3223
	sensation	0.59	0.70	0.64	2862
	accuracy			0.63	6085
	macro avg	0.64	0.64	0.63	6085
	weighted avg	0.64	0.63	0.63	6085

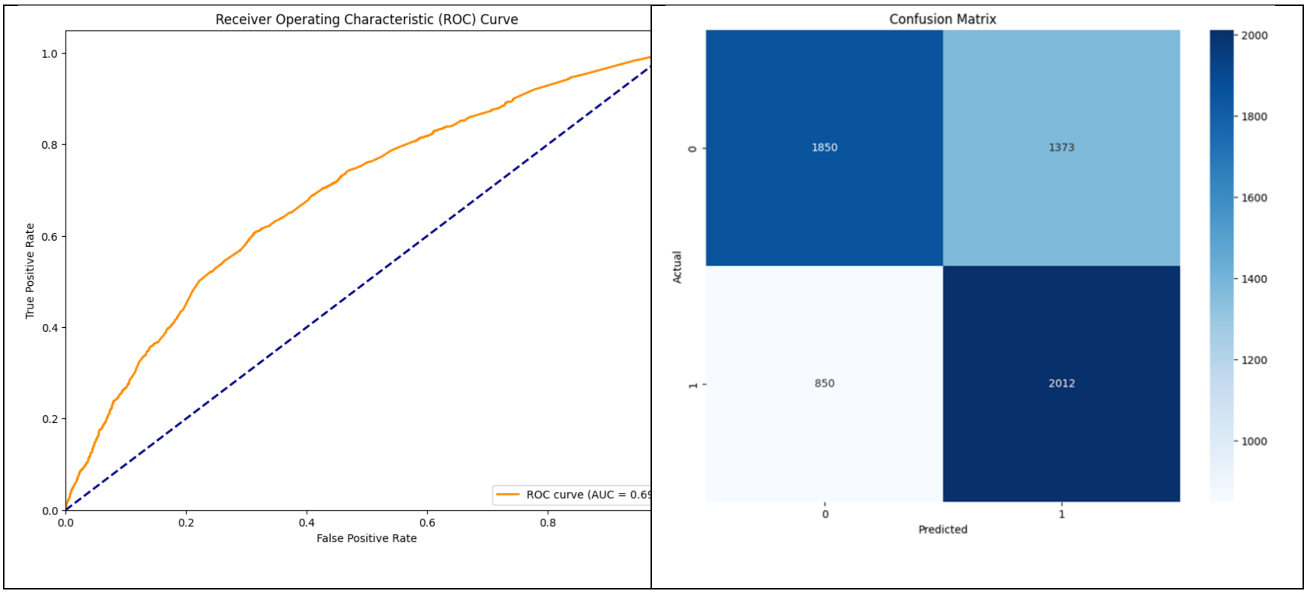


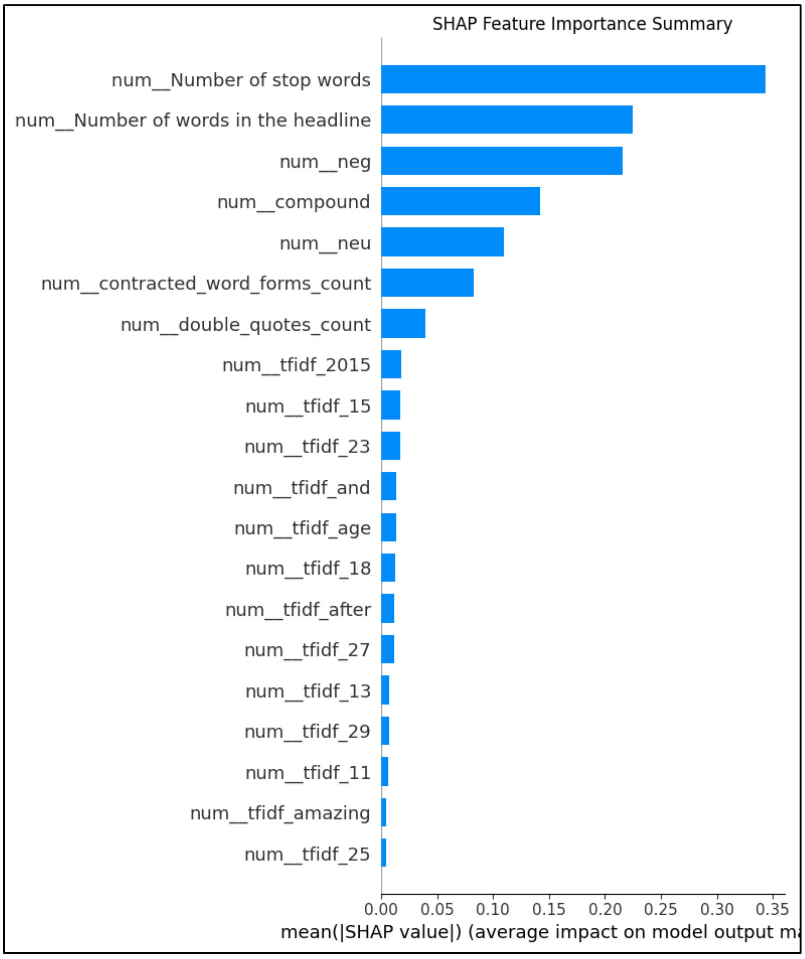
Table 20

<table>
<tr><td rowspan="8">CATBoost on Test set</td><td colspan="5">Best parameters: {'subsample': 0.8, 'random_strength': 0.1, 'learning_rate': 0.1, 'l2_leaf_reg': 3, 'iterations': 100, 'depth': 9, 'colsample_bylevel': 1.0}</td></tr>
<tr><td colspan="5">Validation Set Classification Report:</td></tr>
<tr><td></td><td>precision</td><td>recall</td><td>f1-score</td><td>support</td></tr>
<tr><td>non-sensation</td><td>0.68</td><td>0.68</td><td>0.68</td><td>3224</td></tr>
<tr><td>sensation</td><td>0.64</td><td>0.64</td><td>0.64</td><td>2861</td></tr>
<tr><td>accuracy</td><td></td><td></td><td>0.66</td><td>6085</td></tr>
<tr><td>macro avg</td><td>0.66</td><td>0.66</td><td>0.66</td><td>6085</td></tr>
<tr><td>weighted avg</td><td>0.66</td><td>0.66</td><td>0.66</td><td>6085</td></tr>
</table>

Stage 3: SHAP analysis:

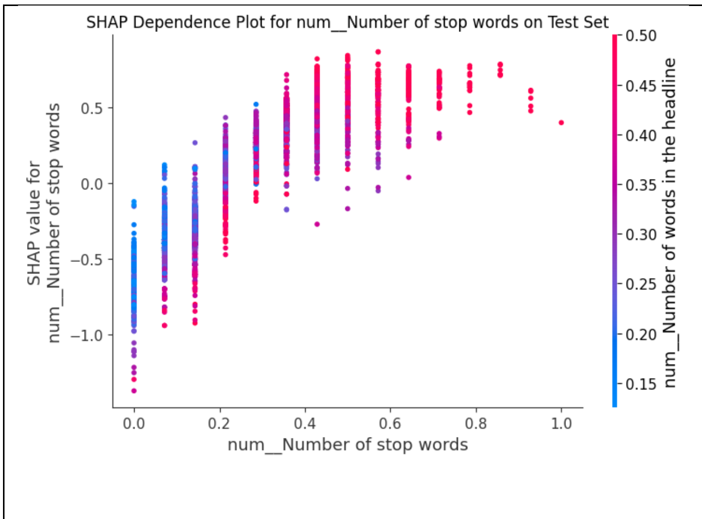
SHAP Feature Importance:

Figure 1



SHAP Dependency:

Figure 2



1. The number of stop words presents a non-linear relationship between with prediction result. Couldn't simply justify if more or less number of stop words can make the prediction more accurate.

2.Breakthrough point appears in the range of 0.2~0.4 (number of stop words). After

	0.2, model tends to predict more headlines as sensational (over 0.0). 3.Obervered positive impact for the model prediction when the number of stop words reaches and over to 0.6.

Figure 3

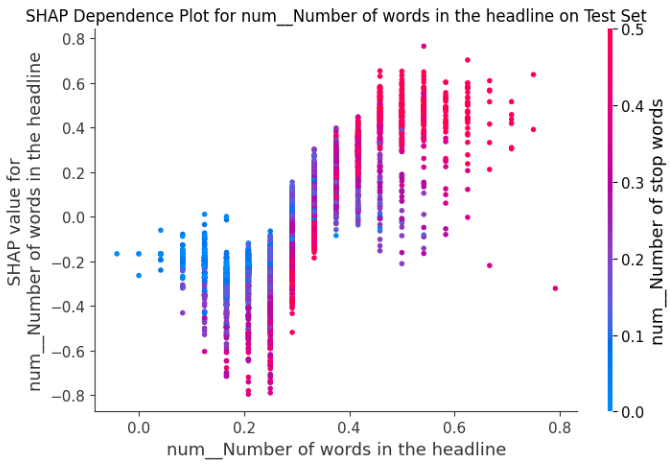
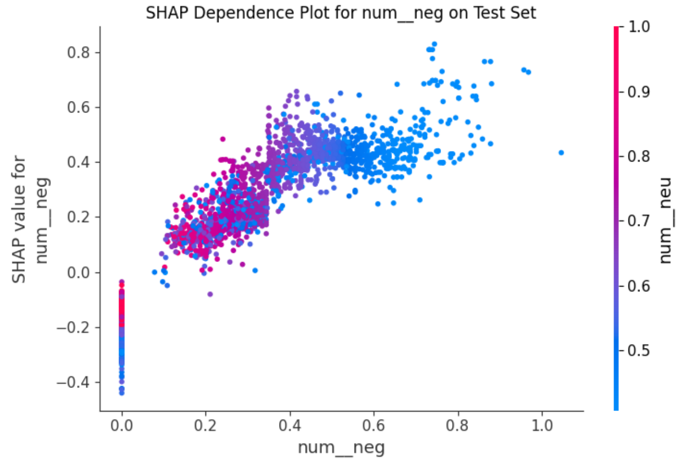
	1.The number of words in the headline presents a non-linear relationship between with prediction result. Couldn’t simply justify if more or less total amount can make the prediction more accurate. 2.Breakthrough point appears in the range of 0.3~0.4 (number of words in the headline). After 0.2, model tends to predict more headlines as sensational (over 0.0). 3.Obervered positive impact for the model prediction when the number of stop words reaches and over to 0.4.

Figure 4

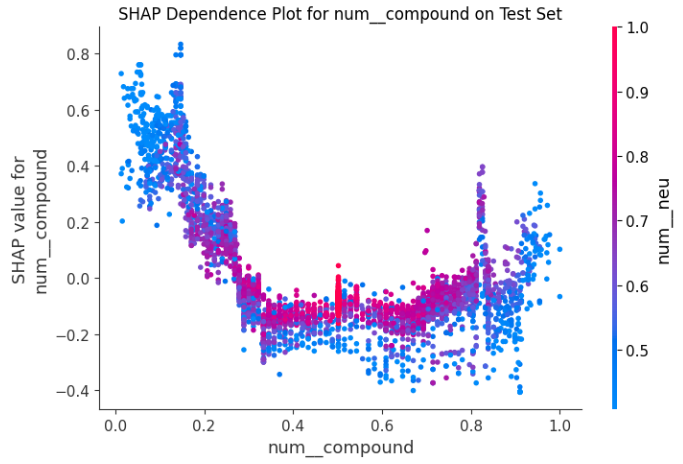


1.Negative sentiment presents a significant positive correlation between features and model prediction results, meaning more negative words results in higher SHAP value.

2.But doesn’t mean a linear relationship between negative sentiment and SHAP value.

3. High variability. Even with a same num_neg, we can observe significant SHAP value differences.

Figure 5



1.Observed non-linear relationships between compound sentiment and prediction results.

2.An inflection point at around num_compound=0.2, indicating that an increase in compound sentiment score could results in a decrease of the prediction value.

3.We can find that when num_com=0.8, high fluctuations in model prediction

	results, meaning the specific language pattern in news headlines.

Figure 6

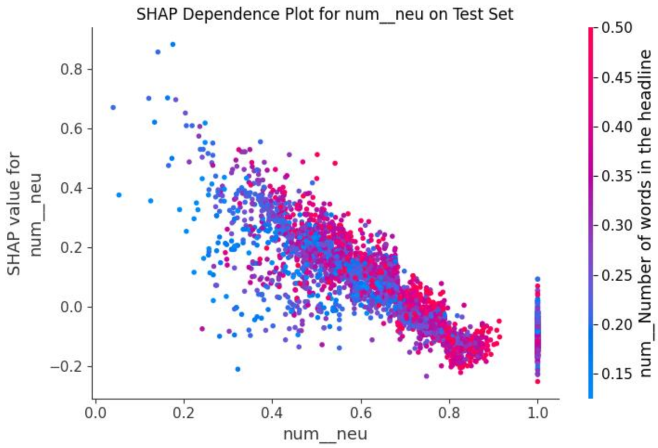
	1.Observed a clear negative relationship between neutral sentiment score and SHAP value. 2.Low SHAP value while with low num_neu, but doesn't mean the existence of linear relationship. Imply the model captured the complex semantic patterns. 3.High variability, especially at low num_neu score area, meaning other features could affect the prediction result.

Figure 7

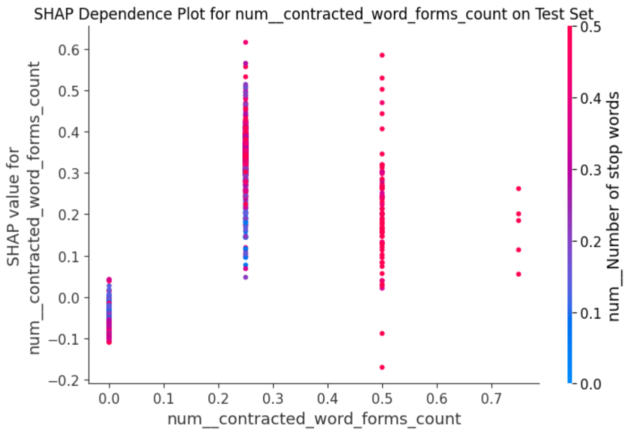
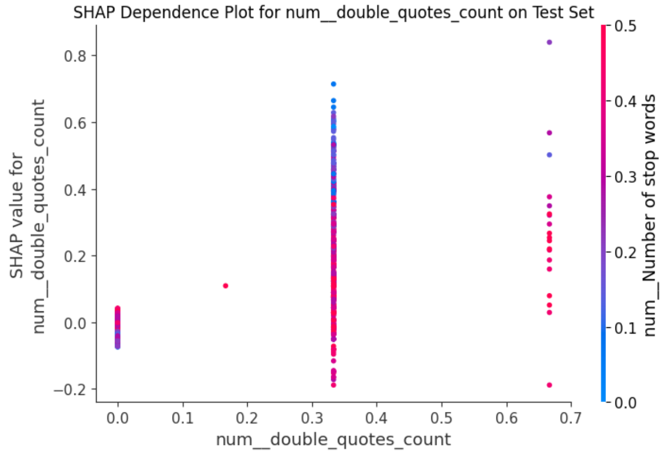
	1.Representing a positive relationship and non-linear relationship between contracted word forms and SHAP value. 2.Number of contracted word from shows a clear discrete characteristic, reflecting the specific pattern of the use in the language. 3.High variability in SHAP value indicates that the complex interaction between features.

Figure 8

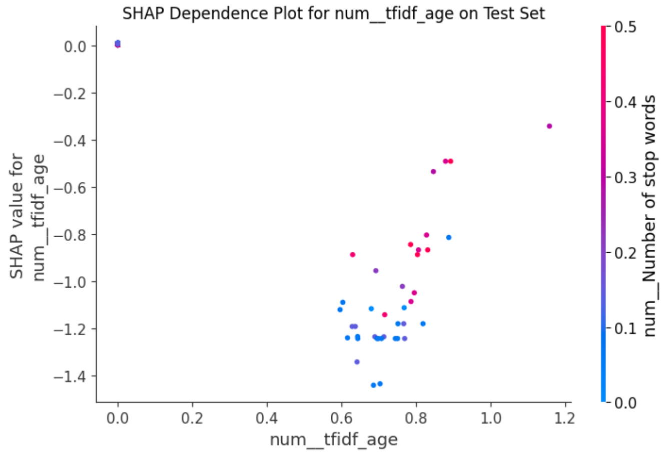


1.Observed threshold effect at double quote count around 0.33 and 0.67, indicating the different expression of headlines structures.

2.Representing highly contextual dependency especially when double quotes count at around 0.33.

3.SHAP values are above 0.0 when the number of double quote count at around 0.67, meaning the positive impact to the prediction result.

Figure 9

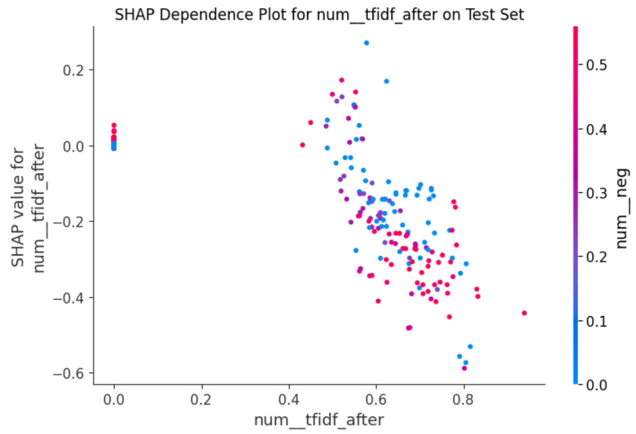


1.Strong negative relationships between TF-IDF_age and SHAP value.

2.The number of TF-IDF_age centralized between 0.6 and 0.8, representing a non-linear relationship and concentration effects that the use of "age" in news headlines.

3.High variability explains the features interaction and the complex contextual dependencies of model prediction.

Figure 10

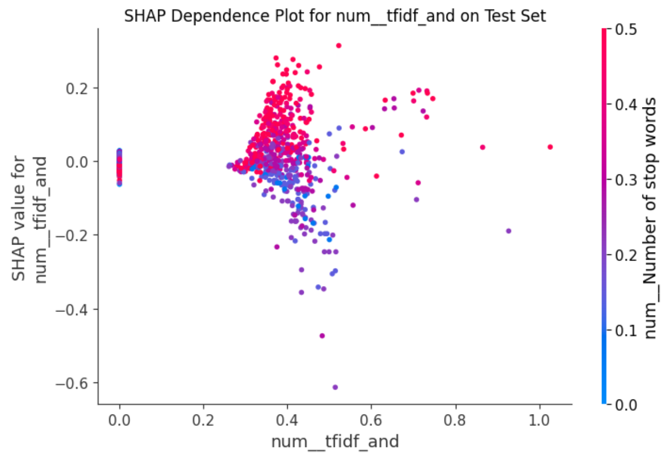


1.Observed significant negative relationship between TF-IDF_after and SHAP value.

2.Representing a non-linear relationship, especially when the value of TF-IDF_after is 0.0, the changes in SHAP value is relatively small and the changes are more significant in the medium to high value region.

3.The discrete pattern represents the contextual dependencies.

Figure 11



1.Observed a complex non-linear relationship between TF-IDF_and and SHAP value.

2.Data points centralize in the range of TF-IDF_and 0.3 to 0.5.

3.The red colour mostly locate above SHAP value 0.0 while TF-IDF_and in the range of 0.3 to 0.5, representing the number of stop words could be a factor that affects the use of “and” and the model prediction.

Chapter 5: Discussion

5.1. Interpretation of Findings

RQ 1. Can we utilize a single linguistic feature to effectively and stability identify the sensational language in news headlines?

Table 21

Feature	Best CV score	Accuracy	f1-score non-sensation	f1-score sensation	f1-score difference in absolute value
Number of words	0.6068	0.60	0.59	0.61[1st]	0.02
Number of stop words	0.6304[1st*]	0.59	0.57	0.61[1st]	0.04
The ratio of stop words to content words	0.6089	0.59	0.59	0.59[2nd]	0.00[1st]
Flesch-Kincaid Readability	0.6089	0.59	0.59	0.59[2nd]	0.00[1st]
Subjectivity and Objectivity	0.6147	0.56	0.52	0.59[2nd]	0.07
Subjectivity	-	-	-	-	-
Objectivity	-	-	-	-	-
Sentiment analysis	0.5854	0.62[3rd]	0.66	0.57[3rd]	0.09
Negative Sentiment	-	-	-	-	-
Neutral Sentiment	-	-	-	-	-
Positive Sentiment	-	-	-	-	-
Compound Sentiment	-	-	-	-	-
Elongated Words	0.3976	0.53	0.69[2nd]	0.00	0.69

Punctuation	0.3597	0.56	0.67[3rd]	0.34	0.33
Currency symbols	-	-	-	-	-
Exclamation marks	-	-	-	-	-
Question marks	-	-	-	-	-
Ellipsis	-	-	-	-	-
Emphasis marks	-	-	-	-	-
Multiple exclamation marks	-	-	-	-	-
Single quotes	-	-	-	-	-
Double quotes	-	-	-	-	-
Contracted word forms	-	-	-	-	-
TF-IDF with Stop words	0.6168[3rd*]	0.66[1st]	0.71[1st]	0.59[2nd]	0.12[2nd]
TF-IDF without Stop words	0.5346	0.61	0.69[2nd]	0.45	0.24
Syntactic 4-grams	0.6233[2nd*]	0.64[2nd]	0.71[1st]	0.55	0.16[3rd]

* [1st], [2nd], and [3rd] indicate the ranking of the feature under different evaluation metrics

1.According to the table, all feature has limited capability to identify sensational news headlines, since the highest f1-score of sensation is only 0.61.

2.We expect a higher score even though the feature The ratio of stop words to content words and the feature Flesch-Kincaid Readability achieved a balance on the f1-score between non-sensational headlines and sensational headlines.

3.In identifying non-sensational headlines, both TF-IDF with Stop words and Syntactic 4-grams perform the best.

4.The feature of Number of Words and Number of Stop words could be the key to identifying sensational news headlines.

5.Compared with TF-IDF without Stop words and TF-IDF with Stop words, stop words represent a more important feature.

We can conclude that the ability to identify sensational news headlines is the breakthrough point to building a robust model. However, we can't identify a valuable feature that mostly contributes to sensational news headlines.

RQ 2. What algorithm demonstrates robust performance in identifying sensational language within news headlines?

1.According to the Tabel 18 and Table 15, both CATBoost and XGBoost performances are similar and robust, yet CATBoost has a higher cross-validation score 0.6697 than XGBoost's 0.6634.

2. XGBoost with Threshold can be chosen if the scenario requires identifying sensational news headlines as much as possible since XGBoost has a higher recall of sensational language.

We can conclude that CATBoost as our most robust model.

RQ 3. Which features most effectively identify sensational language in news headlines among selected features?

1.We can find the Number of stop words is the tier 1 important feature. The Number of words in News Headlines and negative sentiment can be tier 2, negative sentiment, compound sentiment, neutral sentiment, Contracted word forms, and Double quotes belong to tier 3.

2.We don't find any syntactic 4-gram features either in SHAP importance or SHAP dependency plot.

5.2. Implications for Practice

Explainable AI:

Feature Fusion Strategy:

Adapt feature fusion strategy, such as Static Feature Fusion (concatenate features with BERT output embeddings) and Dynamic Feature Fusion (directly combine with key features at the input stage of

BERT) to preserve the BERT's contextual understanding and considering the impact of specific linguistic features on the prediction.

Multi-level Feature Evaluation:

Categorize features into different level (tier1, tier2, tier 3…) to evaluate the dependence of BERT model on different features by layered approach.

Optimization news headlines construction:

Number of stop words and number of words in news headlines is the key features of sensational news headlines. The finding indicates that we should take consideration about the use of number of stop words and control the length of headlines to increase the attraction of news headlines.

Prove negative sentiment indeed increase the attraction of news headlines. Controversially, the use of negative sentiment should be carefully considered as well.

Improvement of recommend system:

The research results show that the performance of TF-IDF with stop words stands out. Content recommend system can be more accurately filter out headline with "sensational feature" headlines to improve audiences' click through rate and engagement. In other hand, it also provides a way of optimization.

Model selection and application:

In the scenario that requires to identify as more sensational news headlines, XGBoost with optimal threshold is the first choice, due to the high recall. However, the robust model is CATBoost that has a better balance between precision and recall.

5.3. Recommendations for Future Research

Table 22

Stage	Future Research
	Use different pre-trained language models.

Dataset Annotation	Specify different identities in the prompt, such as age, gender, profession, expertise, health condition, etc.
	Specify personality such as Myers–Briggs Type Indicator (MBTI).
	Simulate a specific situation as prompt while pre-trained language models annotating (read the news headlines).
	Combine or compare with human annotations.
Feature Extraction	Apply robust part-of-speech tag and NER methods such as pre-trained language model.
	Adapt more linguistic features such as forward-reference, rhetorical device, demonstrative Adjectives, modal verbs, etc.
	Apply unsupervised topic modelling techniques such as k-means or LDA (Latent Dirichlet Allocation) as features.
	Use synonym replacement (word embedding, synonym dictionary) and data augmentation.
	Discuss suitable stop words list with professional in news industry.
Model Implementation	Adjust self-defined optimal threshold and train different models..
	Scale up hyper parameter space.
	Examine the instance of correct predictions and wrong predictions for deep understanding.
	Multimodal and Multilingual Analysis by using photos, videos, social media metrics and metadata.
	Use deep learning models such as Bidirectional Encoder Representations from Transformers (BERT) to achieve better performance and research explainable AI.

Feature Clustering	Apply clustering algorithms (e.g., k-means, hierarchical clustering) to group similar features.
	Use the clustered features to build a simpler surrogate model, such as a decision tree to visualize the key decision paths and how clustered features contribute to predictions.
	Add important features from SHAP to attention input layer to train model.

Chapter 6: Conclusion

6.1. Summary of Findings

- Brief recap of the key findings.

6.2. Conclusions Drawn

- Main conclusions based on the research questions or hypotheses.

We create a MIRUKU sensational news headlines dataset for future research and complete a ground-breaking research by using machine leaning and natural language processing methods in identifying sensational language in news headline. This research illustrates the complexity of human language and emotion since we can’t simply choose a features to effectively decide if the news headline is sensational or not. We combine lexical features, syntactic features, semantic features, readability feature, and stylistic features to obtain a robust model. We verified that negative sentiment is one of the key features to arouse the emotion and capture attention and identify important features in SHAP analysis, paving the way to the research of sensational language in explainable AI.

References

1. Astari, R., Perawironegoro, D., Faturrahman, M.I. and Kistoro, H.C.A., 2023. Indonesian Muslim society's reception of sensation language and invitation to polygamy on social media. Indonesian Journal of Islam and Muslim Societies, 13(2), pp.369-397. Available at: https://doi.org/10.18326/ijims.v13i2.369-397

2. Arbaoui, B., De Swert, K. and van der Brug, W., 2020. Sensationalism in news coverage: A comparative study in 14 television systems. Communication Research, 47(2), pp.299-320. Available at: https://doi.org/10.1177/0093650216663364
3. Arlim, N., Kushadiani, S.K., Riyanto, S., Rodiah, R., Arianty, R., Maukar, M., Al Hakim, S. and Maulana, A.H.A., 2022. Sarcasm Detection in Indonesian Tweets Using Hyperbole Features. In: Proceedings of the 2022 International Conference on Computer, Control, Informatics and Its Applications. IC3INA '22. Association for Computing Machinery, pp.130-134. Available at: https://dl.acm.org/doi/10.1145/3575882.3575908
4. Basera, J., 2023. AN ANALYSIS OF SENSATIONALISM IN NEWS. Journal of the Asiatic Society of Mumbai, XCV(45), pp.168-177. Available at: https://www.researchgate.net/publication/374812735_AN_ANALYSIS_OF_SENSATIONALISM_IN_NEWS
5. Bell, M.A., 2015. The Effects of Sensational Language in News on Memory and Attitudes. Unpublished thesis. Texas State University. Available at: https://hdl.handle.net/10877/5545
6. Bergstra, J. and Bengio, Y., 2012. Random Search for Hyper-Parameter Optimization. The Journal of Machine Learning Research, 13, pp.281-305. Available at: https://dl.acm.org/doi/10.5555/2188385.2188395
7. Bharti, S.K., Vachha, B., Pradhan, R.K., Babu, K.S. and Jena, S.K., 2016. Sarcastic sentiment detection in tweets streamed in real time: a big data approach. Digital Communications and Networks, 2(3), pp.108-121. Available at: https://doi.org/10.1016/j.dcan.2016.06.002
8. Biyani, P., Tsioutsiouliklis, K. and Blackmer, J., 2016. 8 amazing secrets for getting more clicks': detecting clickbaits in news streams using article informality. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence. AAAI'16. AAAI Press, pp.94-100. Available at: https://dl.acm.org/doi/10.5555/3015812.3015827

9. Bourgonje, P., Schneider, J.M. and Rehm, G., 2017. From Clickbait to Fake News Detection: An Approach based on Detecting the Stance of Headlines to Articles. In: Proceedings of the 2017 EMNLP Workshop: Natural Language Processing meets Journalism. Association for Computational Linguistics, pp.84-89. Available at: https://doi.org/10.18653/v1%2FW17-4215
10. Brosco, D.M. and Ionescu, R.T., 2023. A Novel Contrastive Learning Method for Clickbait Detection on RoCliCo: A Romanian Clickbait Corpus of News Articles. In: Findings of the Association for Computational Linguistics: EMNLP 2023. Association for Computational Linguistics, pp.9547-9555. Available at: http://dx.doi.org/10.18653/v1/2023.findings-emnlp.640
11. Brown, D.K., Harlow, S., García-Perdomo, V. and Salaverría, R., 2018. A new sensation? An international exploration of sensationalism and social media recommendations in online news publications. Journalism, 19(11), pp.1497-1516. Available at: https://doi.org/10.1177/1464884916683549
12. Buschmeier, K., Cimiano, P. and Klinger, R., 2014. An Impact Analysis of Features in a Classification Approach to Irony Detection in Product Reviews. In: Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis. Association for Computational Linguistics, pp.42-49. Available at: https://doi.org/10.3115/v1/W14-2608
13. Cano Mora, L., 2009. All or nothing: A semantic analysis of hyperbole. Revista de Lingüística y Lenguas Aplicadas, 4, pp.25-35. Available at: http://dx.doi.org/10.4995/rlyla.2009.731
14. Chakraborty, A., Paranjape, B., Kakarla, S. and Ganguly, N., 2016. Stop Clickbait: Detecting and preventing clickbaits in online news media. In: Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. ASONAM '16. IEEE Press, pp.9-16. Available at: https://dl.acm.org/doi/10.5555/3192424.3192427

15. Davis, H. and McLeod, S.L., 2003. Why humans value sensational news: An evolutionary perspective. Evolution and Human Behavior, 24(3), pp.208-216. Available at: https://doi.org/10.1016/S1090-5138(03)00012-6

16. Dor, D., 2003. On newspaper headlines as relevance optimizers. Journal of Pragmatics, 35(5), pp.695-721. Available at: https://doi.org/10.1016/S0378-2166(02)00134-0

17. Hallman, P., 2016. All and every as quantity superlatives. In: Proceedings of SALT 26. Linguistic Society of America, pp.506-525. Available at: http://dx.doi.org/10.3765/salt.v26i0.3798

18. Hendriks Vettehen, P. and Kleemans, M., 2017. Proving the Obvious? What Sensationalism Contributes to the Time Spent on News Video. Electronic News, 12(2), pp.113-127. Available at: http://dx.doi.org/10.1177/1931243117739947

19. Hendriks Vettehen, P. and Peeters, A., 2008. Explaining Effects of Sensationalism on Liking of Television News Stories: The Role of Emotional Arousal. Communication Research, 35(3), pp.319-338. Available at: http://dx.doi.org/10.1177/0093650208315960

20. Hoffmana, S.J. and Justicz, V., 2016. Automatically quantifying the scientific quality and sensationalism of news records mentioning pandemics: validating a maximum entropy machine-learning model. Journal of Clinical Epidemiology, 75, pp.47-55. Available at: https://doi.org/10.1016/j.jclinepi.2015.12.010

21. Imran, A.A., Shovon, M.S.H. and Mridha, M.F., 2024. BaitBuster-Bangla: A comprehensive dataset for clickbait detection in Bangla with multi-feature and multi-modal analysis. Data in Brief, 53, p.110239. Available at: https://doi.org/10.1016/j.dib.2024.110239

22. Indurthi, V., Oota, S.R., Gupta, M. and Varma, V., 2018. Believe it or not! Identifying bizarre news in online news media. In: Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. CODS-COMAD '18. Association for Computing Machinery, pp.257-264. Available at: https://dl.acm.org/doi/10.1145/3152494.3152524

23. Joy, S.S., Aishi, T.D., Nodi, N.T. and Rasel, A.A., 2023. BanglaClickBERT: Bangla Clickbait Detection from News Headlines using Domain Adaptive BanglaBERT and MLP Techniques. In: Proceedings of the 21st Annual Workshop of the Australasian Language Technology Association. Association for Computational Linguistics, pp.1-10. Available at: https://aclanthology.org/2023.alta-1.1/
24. Karoui, J., Benamara, F., Moriceau, V., Patti, V., Bosco, C. and Aussenac-Gilles, N., 2017. Exploring the Impact of Pragmatic Phenomena on Irony Detection in Tweets: A Multilingual Corpus Study. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers. Association for Computational Linguistics, pp.262-272. Available at: http://dx.doi.org/10.18653/v1/E17-1025
25. Kleemans, M. and Hendriks Vettehen, P., 2009. Sensationalism in television news: A Review. In: R.P. Konig, P.W.M. Nelissen and F.J.M. Huysmans, eds. Meaningful media: Communication research on the social construction of reality. Ubbergen: Tandem Felix, pp.226-243. Available at: https://hdl.handle.net/2066/74717
26. Liu, T., Yu, K., Wang, L., Zhang, X., Zhou, H. and Wu, X., 2022. Clickbait detection on WeChat: A deep model integrating semantic and syntactic information. Knowledge-Based Systems, 245, p.108605. Available at: https://doi.org/10.1016/j.knosys.2022.108605
27. Molek-Kozakowska, K., 2013. Towards a pragma-linguistic framework for the study of sensationalism in news headlines. Discourse & Communication, 7(2), pp.173-197. Available at: https://doi.org/10.1177/1750481312471668
28. Molina, M.D., Sundar, S.S., Rony, M.M.U., Hassan, N., Le, T. and Lee, D., 2021. Does Clickbait Actually Atract More Clicks? Three Clickbait. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. CHI '21. Association for Computing Machinery, pp.1-19. Available at: https://dl.acm.org/doi/10.1145/3411764.3445753

29. Oraby, S., Harrison, V., Reed, L., Hernandez, E., Riloff, E. and Walker, M., 2016. Creating and Characterizing a Diverse Corpus of Sarcasm in Dialogue. In: Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue. Association for Computational Linguistics, pp.31-41. Available at: https://doi.org/10.18653/v1/W16-3604
30. Sangwan, S.R. and Bhatia, M.P.S., 2020. D-BullyRumbler a safety rumble strip to resolve online denigration bullying using a hybrid flter-wrapper approach. Multimedia Systems, 28, pp.1987-2003. Available at: https://doi.org/10.1007/s00530-020-00661-w
31. Singh, V. (1996). News Clickbait Dataset. Kaggle. Available at: https://www.kaggle.com/datasets/vikassingh1996/news-clickbait-dataset [Accessed 29 August 2024].
32. Uribe, R. and Gunter, B., 2007. Are 'Sensational' News Stories More Likely to Trigger Viewers' Emotions than Non-Sensational News Stories: A Content Analysis of British TV News. European Journal of Communication, 22(2), pp.207-228. Available at: http://dx.doi.org/10.1177/0267323107076770
33. Vista, A., 2014. Mass media, the 'sensational message', and metamorphic truths. Telematics and Informatics, 32(2), pp.416-423. Available at: https://doi.org/10.1016/j.tele.2014.05.005
34. Volkova, S., Shaffer, K., Jang, J.Y. and Hodas, N., 2017. Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers). Association for Computational Linguistics, pp.647-653. Available at: https://doi.org/10.18653/v1/P17-2102
35. William, A. and Sari, Y., 2020. CLICK-ID: A novel dataset for Indonesian clickbait headlines. Data in Brief, 32, p.106231. Available at: https://doi.org/10.1016/j.dib.2020.106231
36. Wuy, C., Wuz, F., Qiy, T. and Huang, Y., 2020. Clickbait Detection with Style-aware Title Modeling and Co-attention. In: Proceedings of the 19th Chinese National Conference on

Computational Linguistics. Chinese Information Processing Society of China, pp.1143-1154. Available at: https://aclanthology.org/2020.ccl-1.106/

37. Yang, F., Mukherjee, A. and Dragut, E., 2017. Satirical News Detection and Analysis using Attention Mechanism and Linguistic Features. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics, pp.1979-1989. Available at: https://doi.org/10.18653/v1/D17-1211

38. Zhang, C. and Clough, P.D., 2020. Investigating clickbait in Chinese social media: A study of WeChat. Online Social Networks and Media, 19, p.100095. Available at: https://doi.org/10.1016/j.osnem.2020.100095

39. Zheng, J., Yu, K. and Wu, X., 2021. A deep model based on Lure and Similarity for Adaptive Clickbait Detection. Knowledge-Based Systems, 214, p.106714. Available at: http://dx.doi.org/10.1016/j.knosys.2020.106714