

Probabilistic Safety Assessment & Management 16 (PSAM 16),
Honolulu, Hawaii, United State,
June 26 – July 1, 2022.



Development of Classification Model for Public Perception of Nuclear Energy in Social Media Platform using Machine Learning: Facebook Platform in Thailand

Wasin VECHGAMA

Ph.D. Student,

Korea Atomic Energy Research Institute (KAERI School),
University of Science and Technology (UST).
Email: wasin@kaeri.re.kr, wasin@ust.ac.kr

Nuclear Engineer,

Nuclear Technology Research and Development Center,
Thailand Institute of Nuclear Technology (Public Organization).

Email: wasin@tint.or.th, wasinvechgama@gmail.com

Team



Wasin Vechgama*

- Thailand Institute of Nuclear Technology (Public Organization)
- Korea Atomic Energy Research Institute
- University of Science and Technology

Watcha Sasawattakul

- Department of Computer Engineering, Faculty of Engineering, Chulalongkorn University

Kampanart Silva

- National Energy Technology Center, National Science and Technology Development Agency

Jaehyun Cho

- Korea Atomic Energy Research Institute
- University of Science and Technology

<https://medium.com/goodai-news/understanding-the-public-perception-of-ai-a14b0e6b6154>



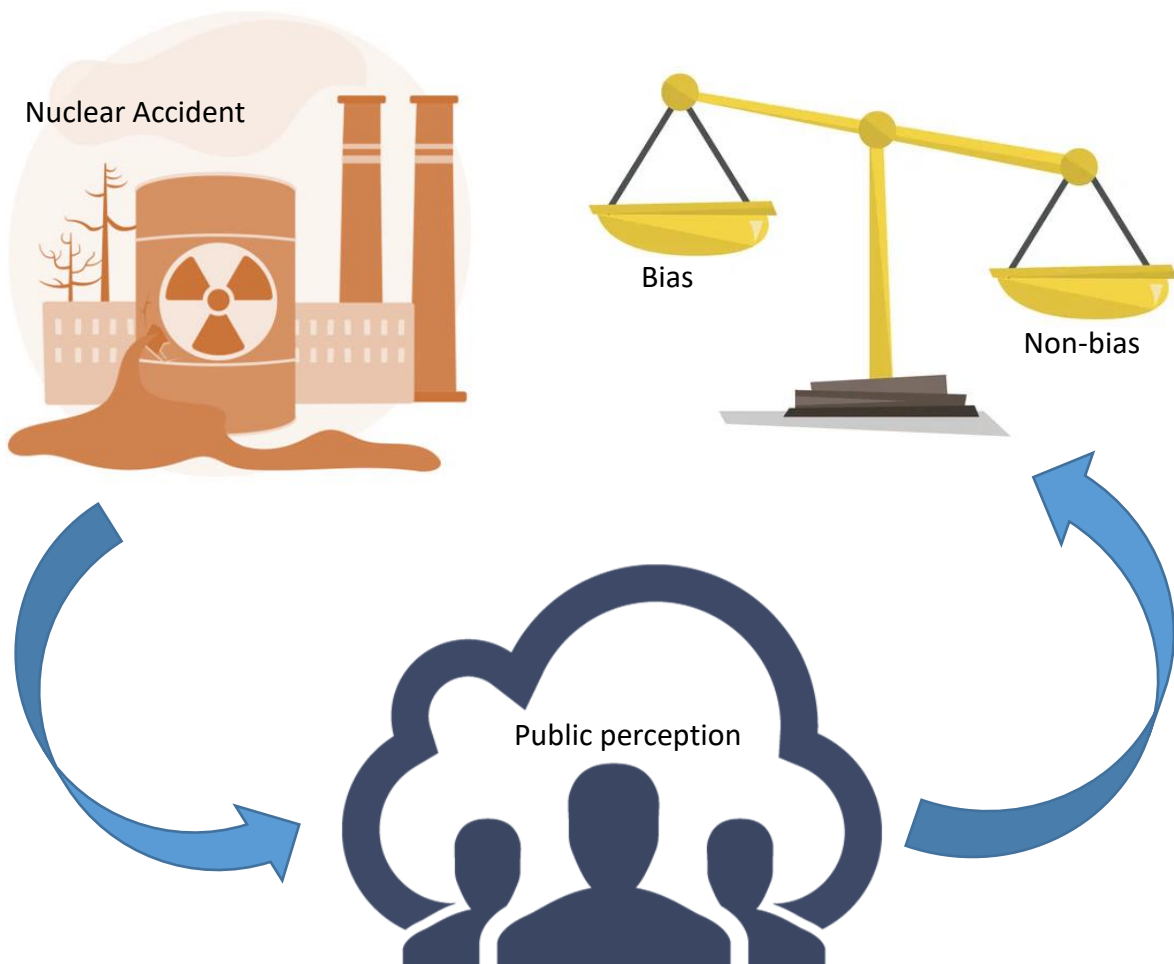
Outline

- ❑ INTRODUCTION AND BACKGROUND
- ❑ METHODOLOGY
- ❑ RESULTS AND DISCUSSION
- ❑ CONCLUSION

<https://letstalkscience.ca/educational-resources/stem-in-context/what-are-pros-and-cons-nuclear-energy>

INTRODUCTION AND BACKGROUND

Nuclear Energy Public Perception



NPP severe accident was recognized as the main input affecting public perception.

Public perception was processed and expressed from each individual thought from many ways.

Basic expressions of people can be learned as output to judge that it is bias and non-bias expressions from the big data that were received.

The one of important things affecting expression output of nuclear energy public perception is the **source of data**.

Nuclear Energy Public Perception in Thailand

Thailand explored the people's opinions about nuclear energy in Thailand using questionnaires in the past.

Pro

- Statistical results might help reflect some aspects of public perception toward nuclear power in Thailand.

Cons

- The survey is time-consuming and not cost-efficient.
- A number of questions to answer affected people might feel troublesome and would not utterly express their feelings and opinions.

- Embarrassed
- Time-consuming
- Biased opinion

Audience

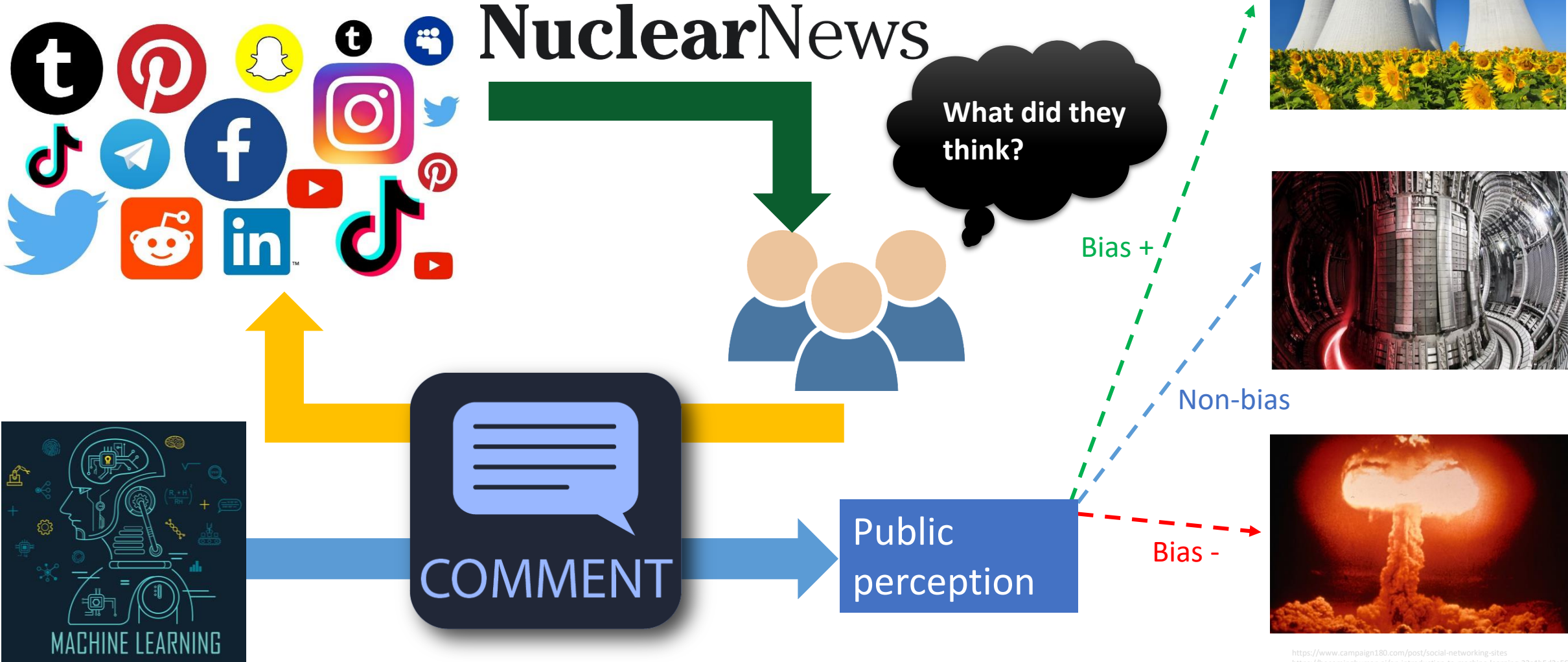


- Time-consuming
- High cost-investment
- Fewer samples

Host

<https://www.peterlugtig.com/post/studying-mode-effects/>

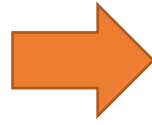
Nuclear Energy Public Perception in Thailand



<https://www.campaign180.com/post/social-networking-sites>
<https://becominghuman.ai/an-introduction-to-machine-learning-33a1b5d3a560>
<https://www.ans.org/pubs/magazines/nn/>
<http://tpsalethiaf.blogspot.com/2017/12/replying-back-2-my-comments-5.html>
<https://www.conservationmagazine.org/2014/12/is-nuclear-power-key-to-biodiversity/>
<https://www.popscl.com/survive-nuclear-radiation/>
<https://www.bbc.com/news/science-environment-60312633>
<https://pixabay.com/vectors/people-icon-three-circles-simple-1085695/>

Nuclear Energy Public Perception in Thailand

Objective



- To develop classification models for public perception of the nuclear energy of Thai people on **Facebook** platform using machine learning techniques.
- To use the classification models to understand Thai public perception on nuclear energy news from **Facebook** including
 - **Non-biased opinions (neutral opinion)**
 - **Bias opinions (both positive and negative opinions)**

https://portlandenglish.edu/learn-english-blog/newsletter_02-17/facebook-2/
<https://www.canstockphoto.com/illustration/computers.html>
<https://www.vectorstock.com/royalty-free-vector/colorful-human-head-with-brain-icon-graphic-vector-9416297>

Nuclear Energy Public Perception in Thailand

Scope

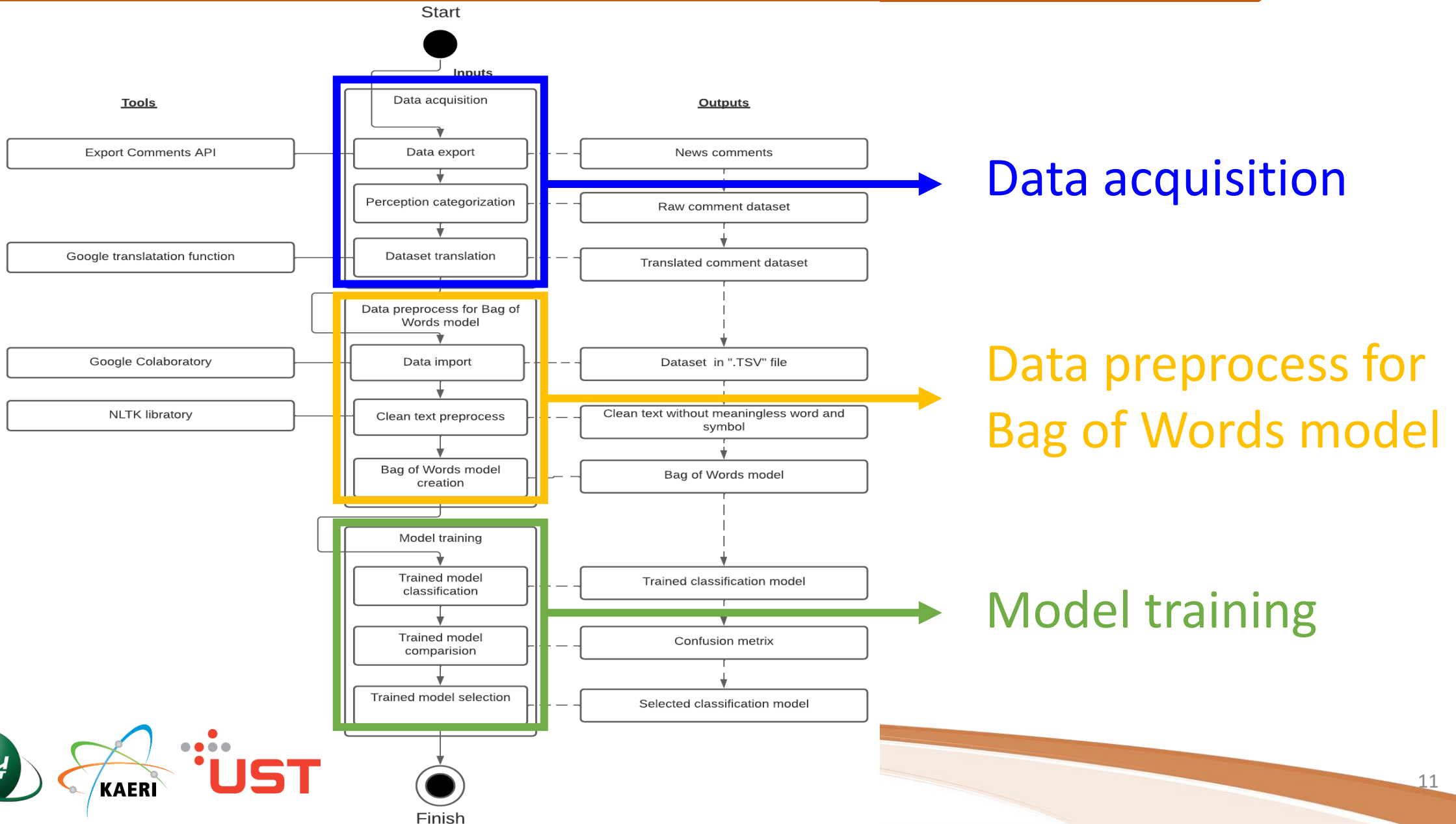
- **Facebook** platform is the main focus of this work in the starting step because Facebook was the most popular in Thailand with the highest number of Thai people who have the Facebook accounts when compared with other social media platforms.
- **All comments** from selected Facebook **news related to the update of nuclear power plant technology in Thai language** were learned in the machine learning process to compare the ability of the classification models and select the proper model to understand public perception.



https://www.facebook.com/story.php?story_fbid=3012822545605432&id=1526071940947174
<https://www.facebook.com/XinhuaNewsAgency.th/posts/3129622220587112>

METHODOLOGY

Classification Model Steps for Learning Public Perception



Data Acquisition

Data export:

- Selected international news related to the update of nuclear power plant technology in the **Thai language** was collected from the Facebook using Export Comments in which a higher number of comments on the news was the priority.

Perception categorization:

- Received Thai comments were analyzed for classifying that is neutral opinions or biased opinions by **the author's evaluation using the criteria below**.
 - Non-related meaningless sentences, links, figures, and symbols were screened out.
 - Comments in headlines that **do not express sentiments** were cut off.
 - Remaining comments related to the headlines were discussed by the authors and were **classified into neutral opinions or biased opinions**.

Dataset translation:

- Thai comment inputs were translated into the English language** using the Google function to match English algorithms for model training.

Data Preprocess for Bag of Words Model

Data import:

- Translated data was imported to Google Colaboratory to write and execute **arbitrary python code** through the web browser machine learning suitably.

Clean text preprocess:

- The sentences of the dataset that have **no meaning would be removed** and use lower characteristics using the function from the **NLTK library**.
- Poster Stemmer library** was applied to **generalize the wordings to the roots** of their words.

Bag of Words model creation:

- All root word datasets became input to **create the bag of the word and were transformed into numeric** in order to create the word representation in a matrix diagram with individual clean words.

Model Training

Trained model classification:

Example

"I am very excited for nuclear technology"

Dataset

excit fir nuclear technolog

Preprocessing sentence to word dataset

Training

Naive Bayes

Classification model

Answer

- **Naïve Bayes classifier** is a statistical classifier based on the posterior hypothesis using the probability data for the large dataset.
- The classification task is to provide the dataset of sample data to the classes using **supervised learning algorithms**.
- The number of the dataset were **separated into two groups for train data and test data** varying the proportion of them **to classify neutral opinion and bias opinion**.

Model Training

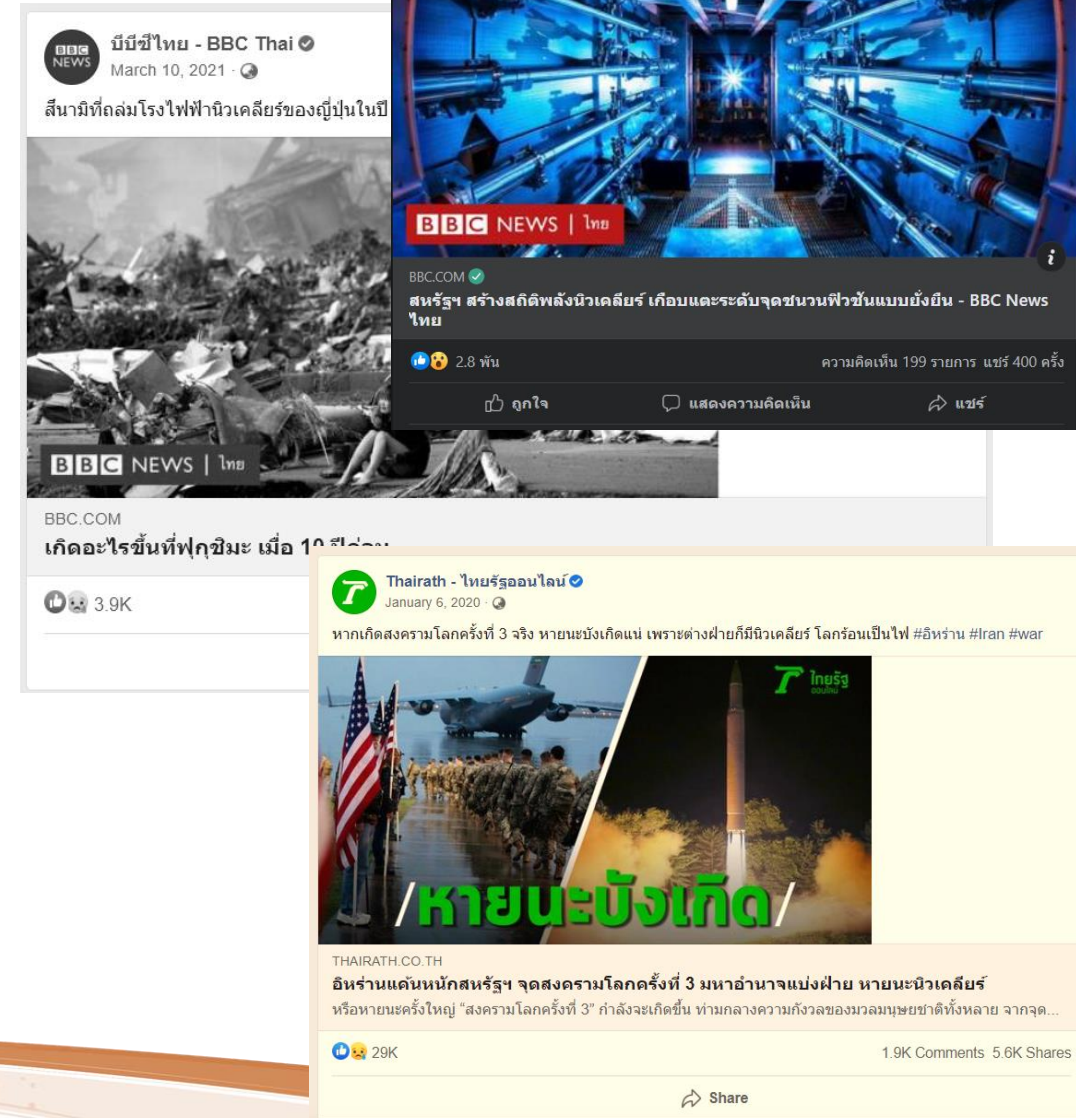
- The **statistical values of accuracy, precision, recall, and F1-score** were calculated to compare the ability of classification models.
- The **highest score** of the statistical values would be considered for selecting the classification model.

Confusion Matrix		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$		
Precision	$\frac{TP}{TP + FP}$		
Recall	$\frac{TP}{TP + FN}$		
F1-score	$\frac{2 \times Precision \times Recall}{Precision + Recall}$		

RESULTS AND DISCUSSION

Data Acquisition

- Nuclear energy news in the Thai language on the Facebook platform in Thailand included the three categories namely;
 - News related to the update of nuclear power plant technology
 - News related to the nuclear accidents and radioactive contamination
 - News related to nuclear weapons and wars
- To avoid having inherent **biases as much as possible** for learning the understanding of the public perception, **only the news related to the update of nuclear power plant technology** was selected to investigate first.



https://www.facebook.com/story.php?story_fbid=3012822545605432&id=1526071940947174
https://www.facebook.com/story.php?story_fbid=2893495587538129&id=1526071940947174&m_entstream_source=timeline
<https://www.facebook.com/thairath/posts/10158872904747439>

News related to the Update of Nuclear Power Plant Technology

Facebook news headlines in translated English	News agency	Country of origin of the news agency	News date	Number of comments	Number of total individual clean words
A new world record, China's artificial sun lasts 1,056 seconds.	China Xinhua New	China	2 Jan 2022	151	700
China contains fuel at the nuclear power plant of the 2 nd place of Hualong-1.	China Xinhua New	China	8 Nov 2021	39	250
China began to build a small commercial nuclear reactor in Hainan.	China Xinhua New	China	14 July 2021	101	550
China's first nuclear power plant to use Hualong-1 technology, a Chinese-developed third-generation reactor design started filling fuel yesterday.	China Xinhua New	China	5 Sep 2020	89	450
Chinese engineers revealed China's first floating nuclear power plant that may work even in extreme weather events.	TNN World	Thailand	16 Dec 2021	96	450
The U.S. breaks a nuclear power record that almost reaches the point of igniting a sustained fusion reaction.	BBC Thai	The UK	20 Aug 2021	177	600
Total				653	1800

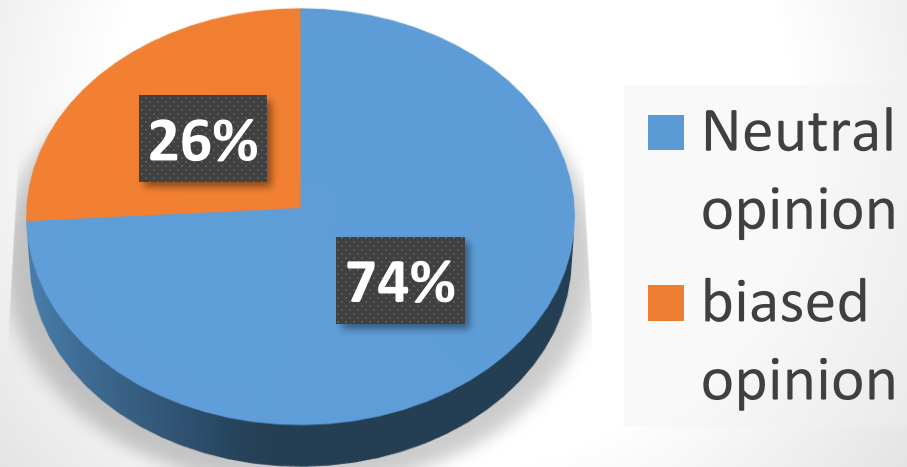
Data Acquisition

- In [2020-2022](#), there were the main three news agencies on the Facebook that concentrated on the news related to the update of nuclear power plant technology in Thailand including
 - [China Xinhua New](#)
 - TNN World
 - BBC Thai
- Due to the [unpopularity of nuclear energy in Thailand in 2020-2022](#), there were only people around tens to hundreds who were interested and commented on each Facebook news post.

Data Acquisition

- All comments were classified into **neutral opinions** or **biased opinions** by the author's judgments, before translating into the English language for the **supervised learning process** in the English algorithm.
- Most Thai people who are interested in nuclear energy on Facebook have neutral opinions about nuclear energy **(74%)**.
- Thus, the results lead to the **opportunity to convince the Thai people on Facebook who have a neutral opinion** to understand and have the proper nuclear perception to increase the nuclear public acceptance in Thailand.

Evaluation of author's judgments for supervised learning process



Data Preprocess for Bag of Words Model

- All Thai comments translated into English were to provide **natural language understanding**.
- Public perception of nuclear energy of Thai people on the Facebook platform mostly came from the **Chinese news** related to the update of nuclear power plant technology due to a lot of media from China Xinhua News in Thailand.



Model Training

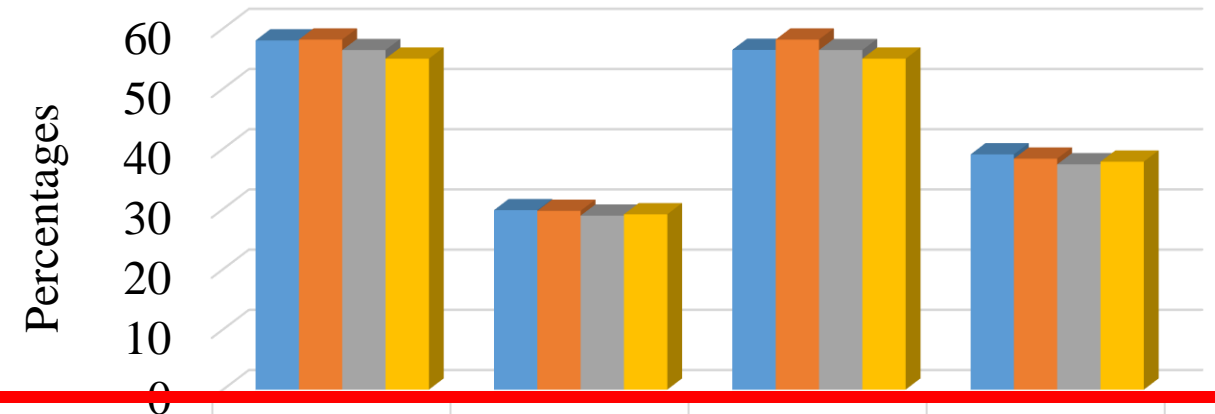
- All comments from news with their natural language were extracted as individual clean words for the training process (1800 words).
- There were four case studies of the training model varying in the **proportion of train data and test data**.
- Train data were used to generate classification models using **Naïve Bays** model.
- The classification models were validated by test data to **suggest appropriate models using the statistical values**.

Case	% Train data	% Test data
1	80	20
2	70	30
3	60	40
4	50	50

Model Training

- Since the results of the comparison of accuracies tended to the close values, **all of the classification models have the potential to use** for the predictions of the public perception of nuclear energy of the Thai people.
- When comparing the accuracy and F1-value (the balance between the precision and the recall) of all classification models, it was found that the classification model in **case 1 has the highest accuracy and F1-value.**

Comparison of statistical values of classification models



	% Accuracy	% Precision	% Recall	% F1_score
Case 1	58	30	56	39
Case 2	58	30	58	38
Case 3	56	29	56	37
Case 4	55	29	55	38

Model Training

- However, the accuracy of the classification model in case 1 was around **only 58%** of the estimation results.
- It seems that the accuracy of the classification model in case 1 was **not quite high enough** to predict the other database.
- *Thus, it was a crucial point for us to look back to the data preprocessing to analyze the problems to improve the model accuracy properly for a future task.*
 - *PB1: The error from the authors' judgment of sentiments in Thai language comments before translating into English for learning in the English algorithm.*
 - *PB2: The insufficient datasets for the learning process.*
 - *PB3: The effect of selected supervised training model.*

CONCLUSIONS

Conclusions

- This study intended to develop classification models to understand the public perception of the nuclear energy of Thai people on Facebook using machine learning.
- Neutral opinion and biased opinion were classified using the Thai comments from news related to the update of the nuclear power plant technology on Facebook.
- The 74% of the people in selected Facebook news have neutral opinions about nuclear energy.
- Thus, it is the opportunity to convince the people who have neutral opinions to understand the proper and precise nuclear perception to support the proper direction of the nuclear energy plan of Thailand as well as an NPP foundation on the public side for other countries not having NPPs.

Conclusions

- In the study, the maximum accuracy of the classification model in case 1 was **around 58%** of the author's judgment results.
- Although the accuracy of the classification model is **quite not enough** in the primary work, it helps us increase more concern the data preprocess of the supervised data to improve the model accuracy.

Wasin VECHGAMA

Ph.D. Student, KAERI School-UST

Email: wasin@kaeri.re.kr, wasin@ust.ac.kr

Nuclear Engineer, TINT

Email: wasin@tint.or.th, wasinvechgama@gmail.com

Q&A

