

# Thai Fake News Detection App

จิรวุฒิ พรหมศรี  
ID:6420421007

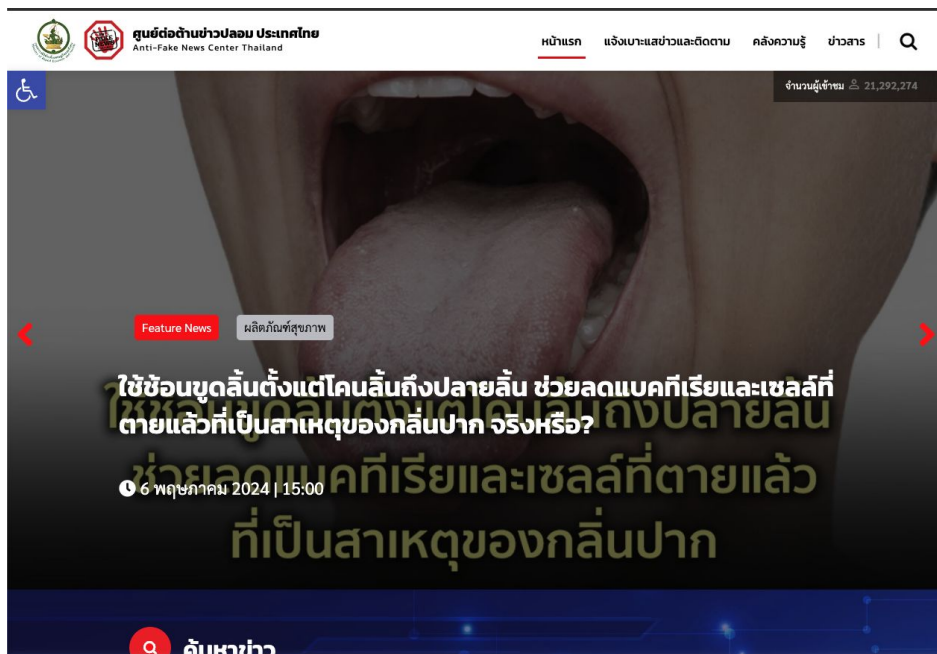
# Intro

ในปัจจุบัน ข่าวออนไลน์เป็นสิ่งที่แพร่กระจายได้อย่างรวดเร็ว และมีประสิทธิภาพต่อผู้ใช้งานเป็นจำนวนมาก แต่ในบางข่าวก็สามารถเป็นได้ทั้งจริงและเท็จ ซึ่งกว่าที่จะรู้ว่าเป็นจริงเป็นเท็จ อาจจะใช้เวลาหลายชั่วโมง เพราะเราไม่สามารถหาต้นตอของข่าว แต่หากว่าเรามี application ที่คอยตรวจสอบ ข่าวได้ว่าจริงหรือเท็จ ก็สามารถช่วยกรอกข้อมูลที่จะมาถึงผู้บริโภคได้อย่างดีขึ้น โดย machine learning



# Dataset Description

- ข้อมูลข่าวจาก <https://www.antifakenewscenter.com/>



# EDA

- split data Train:Test = 80:20
- data datetime : 04/2567-05/2567
- Text size max : 100 character

Class	No. of Text
True	100
Fake	100



สภาวะเอลนีโญส่งผลให้อุณหภูมิโลก และน้ำทะเลสูงขึ้นกว่าเดิม จริงหรือ?

✓ ข่าวจริง

11 มีนาคม 2567 | 16:30 น.



ค่าฝุ่น PM 2.5 ในภาคเหนือสูง เพราะเกิดจากหมอกควันข้ามแดน จริงหรือ?

✓ ข่าวจริง

11 มีนาคม 2567 | 15:00 น.



หากผู้ประกันตนมาตรา 39 ขาดส่งเงินสมทบ 3 เดือน จำเป็นต้องตัดสิทธิประกันสังคม จริงหรือ?

✓ ข่าวจริง

11 มีนาคม 2567 | 11:30 น.



ข่าวปลอม อย่าแชร์! 4 วิธีการปฐมพยาบาลเมื่อสำลักอาหาร ตกหมอน ขาเป็นตะคริว และขาชา

❌ ข่าวปลอม มลพิษทางสุขภาพ

15 พฤษภาคม 2567 | 16:30 น.



ข่าวปลอม อย่าแชร์! กระทรวงคมนาคมประกาศยกเลิกโครงการแลนด์บริดจ์

❌ ข่าวปลอม นโยบายรัฐบาล-ข่าวสาร

15 พฤษภาคม 2567 | 16:00 น.



ข่าวปลอม อย่าแชร์! สามารถรับอุปกรณ์เด็กได้ ผ่านกลุ่มรับอุปกรณ์เด็ก

❌ ข่าวปลอม นโยบายรัฐบาล-ข่าวสาร

15 พฤษภาคม 2567 | 15:00 น.

# Process

## 1. แบ่งคำโดยใช้ pythainlp

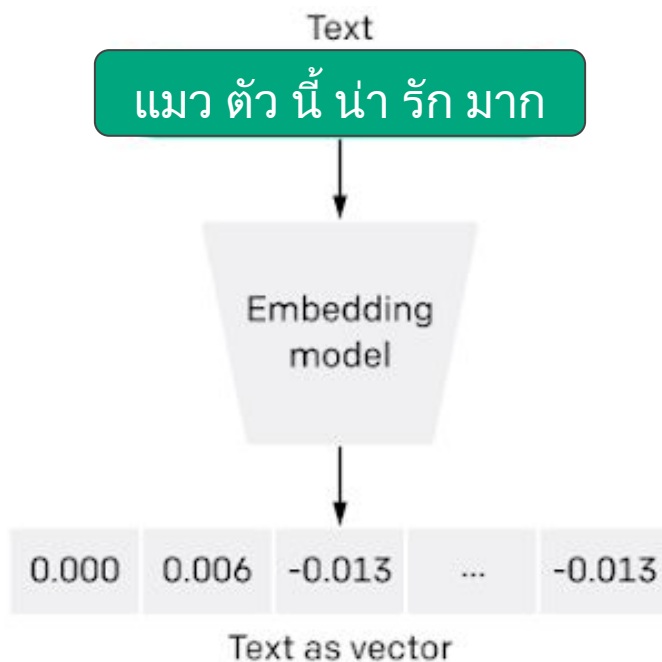
แมวตัวนี้น่ารักมาก



แมว ตัว นี น่ า รัก มาก

# Process

## 2. เปลี่ยน text เป็น vector



# Compare Model

- Logistic Regression
- Decision Tree Classification
- Gradient Boosting Classifier
- Random Forest Classifier
- MultinomialNB



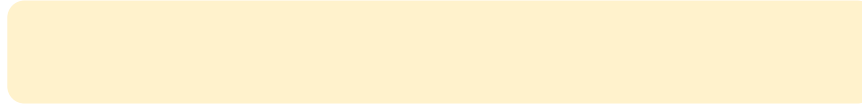
# Compawre Model

**Logistic Regression**



73 %

**Decision Tree Classification**



69 %

**Gradient Boosting Classifier**



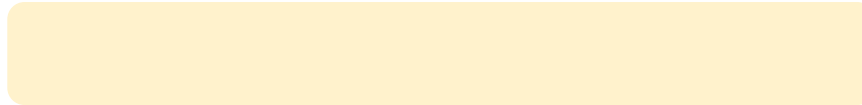
76 %

**Random Forest Classifier**



80 %

**MultinomiaNB**



69 %



# Demo

## Fake News Detection

Enter news text here...

Check



# Improvement in Future

- รองรับ character ให้มากขึ้น
- สามารถตรวจสอบใน twtiter ได้
- ใช้ neural networks RNN ในการ classification ให้ accuracy สูงขึ้น





Thank you for watching