

The background features abstract, translucent geometric shapes in shades of blue, purple, and pink, floating in the upper right quadrant.

Sentiment Analysis

Final Project

DADS6005 : Data Streaming and Real Time Analysis

Contents

Q1

Project
Description

Q2

Diagram

Q3

Modeling Process

- Data Source
- Data Management
- Real Time Process
- Analytics Result

Q4

Insights and
Conclusions

Q1 : Project Discription

This project is about analyzing social media comments on live Streaming of HEARTROCKER Channel and predicting trends with sentiment classification

Q2 : Diagram



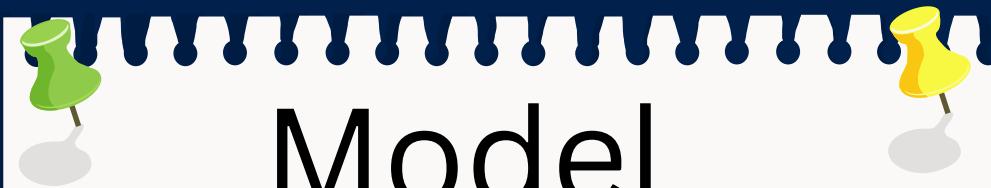
Hive MQ



Pre-Train
Sentiment model



Sentiment model

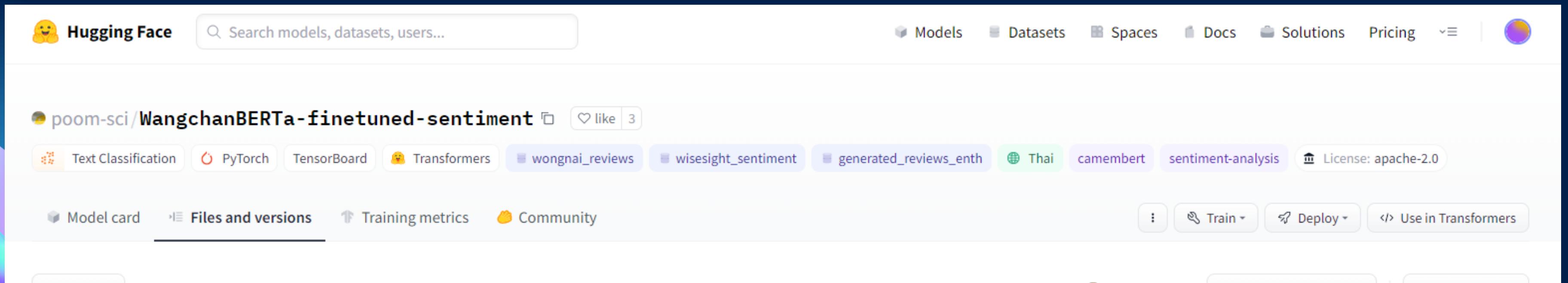


- This project will be using library from WangchanBERTa-finetuned-sentiment to sentimental analysis.

WangchanBERTa: Pre-trained Thai Language Model

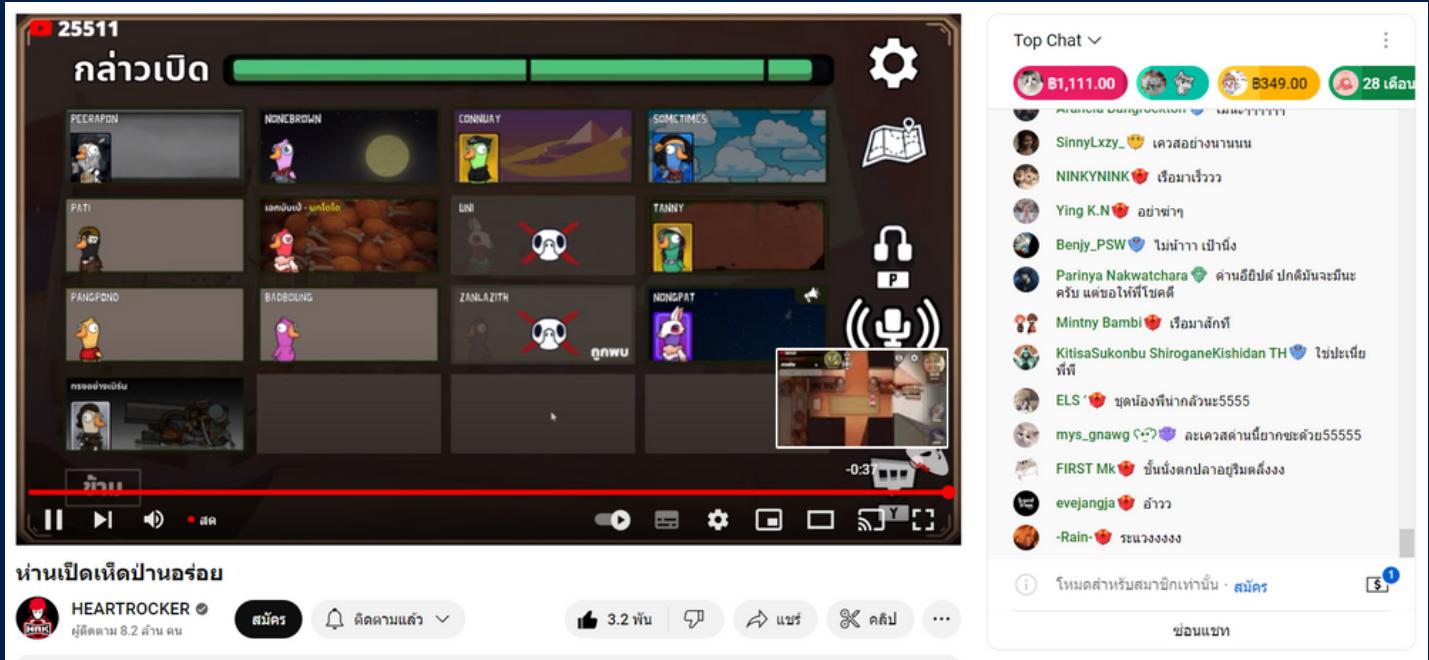
โมเดลภาษาสำหรับงานประมวลผล และการเข้าใจภาษาไทย

สถาบันวิจัยปัญญาประดิษฐ์ประเทศไทย (Thailand Artificial Intelligence Research Institute) ได้ทำการเทรนโมเดลภาษา (language model) บนชุดข้อมูลในภาษาไทยที่ได้จากแหล่งต่างๆ เช่น ข่าว, วิกิพีเดีย, ข้อความในโซเชียลมีเดีย และข้อมูลที่ได้จากการ crawl เว็บไซต์ในอินเทอร์เน็ต ซึ่งมีขนาดข้อมูลรวม 78.5 GB และได้วัดประสิทธิภาพของโมเดลภาษาที่ finetune แล้ว ได้ผลคะแนน micro-averaged F1 score สูงที่สุดบน 5 ชุดข้อมูล จากห้องทดลอง 6 ชุดข้อมูล โดยเป็นชุดข้อมูลทดสอบในโจทย์การจำแนกข้อความ (text classification) และการจำแนกคำ (token classification) เมื่อเทียบกับ baseline model และโมเดลภาษาแบบหลายภาษา (multilingual language model) ที่มีอยู่ในปัจจุบัน (mBERT และ XLMR)



The screenshot shows the Hugging Face Model Card for the 'WangchanBERTa-finetuned-sentiment' model. At the top, there's a navigation bar with icons for Hugging Face, search, models, datasets, spaces, docs, solutions, pricing, and a user profile. Below the search bar, the model name 'poom-sci/WangchanBERTa-finetuned-sentiment' is displayed along with a 'like' button showing 3 likes. A 'Model card' tab is active. Underneath, there are several tags: Text Classification, PyTorch, TensorBoard, Transformers, wongnai_reviews, wiselight_sentiment, generated_reviews_enth, Thai, camembert, sentiment-analysis, and License: apache-2.0. At the bottom, there are buttons for Files and versions, Training metrics, Community, and various deployment and integration options like Train, Deploy, and Use in Transformers.

Q3 : Data Collection and Modeling Process



<https://www.youtube.com/watch?v=EBcvHK2Moys>



Pre-Train Sentiment model



Negative



Neutral

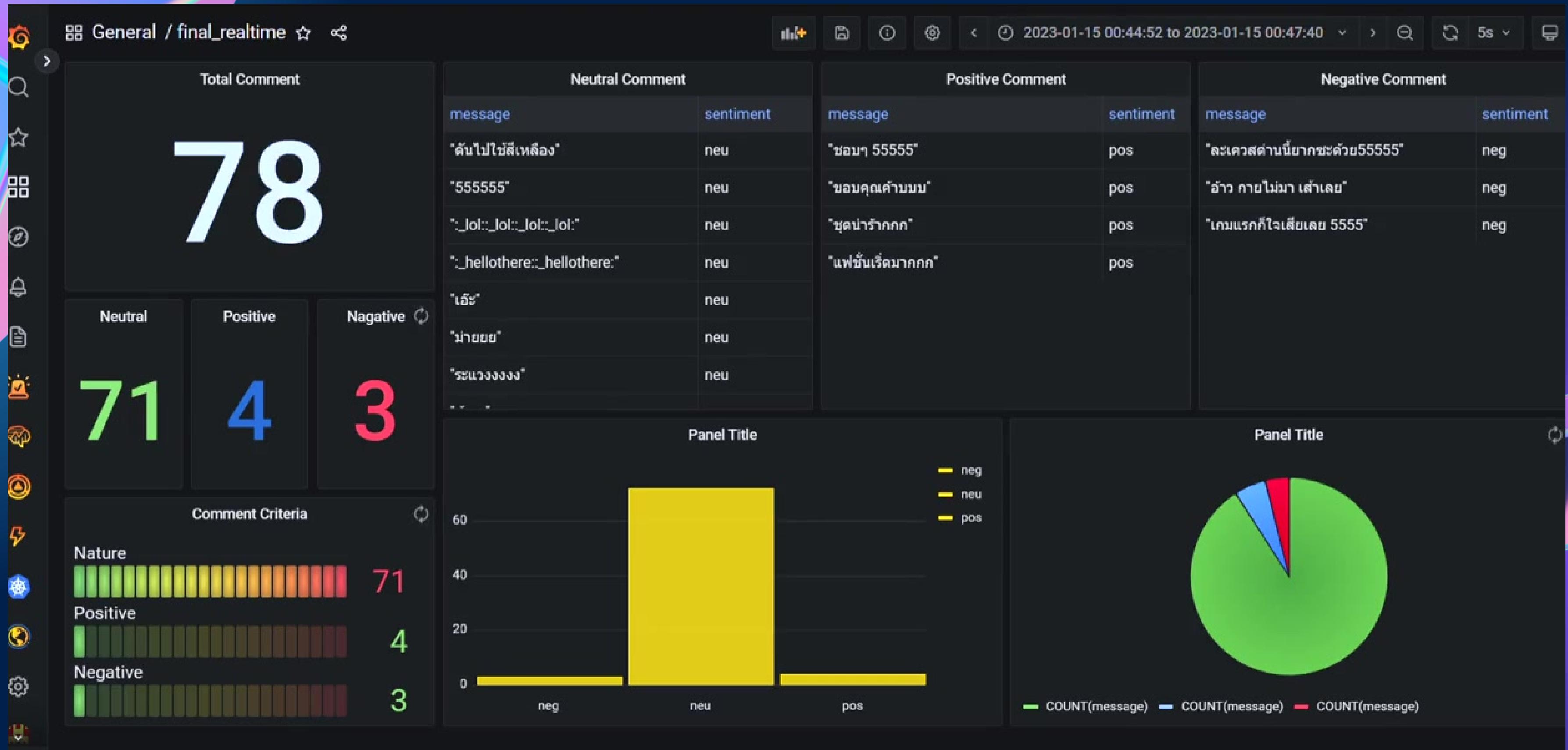


Positive

Source Code : publish

Source Code : subscribe

Analytics Result Visualization with Grafana



Q4 : Conclusions



Real-Time Dashboard Show

1. Total Comment
2. Sentiment Analysis
 - o Positive Comments
 - o Neutral Comments
 - o Negative Comments
3. Comment Text Group by Sentiment Analysis
4. Sentiment Analysis Proportion



Discussion

- Before training model, the model has not high accuracy.
- customer's opinion or recommendations are a treasure trove of information that can be analyzed and processed to assess customer sentiment. So that the webpage can plan and improve for future work.

Our Team



Chokchai Kenpho
6410422004



Noppol Anakpluek
6410422009



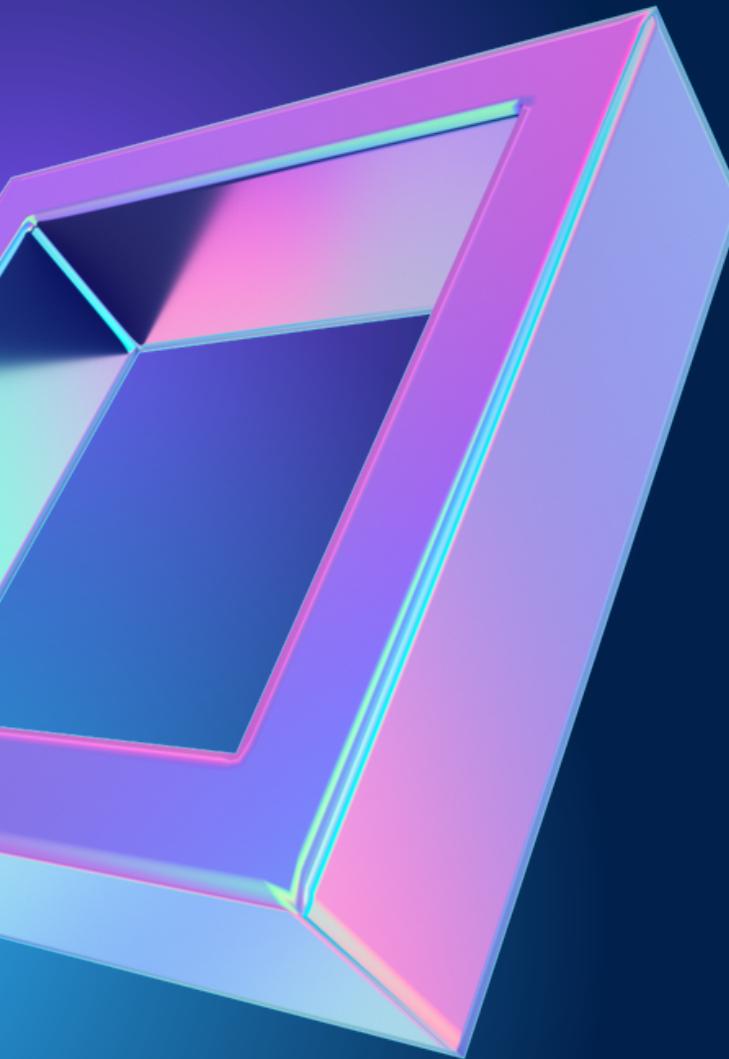
Supanan Sukhamta
6410422020



Natthaporn Wattanakul
6410422026



Watcharakorn Pasanta
6420422006



Thank You