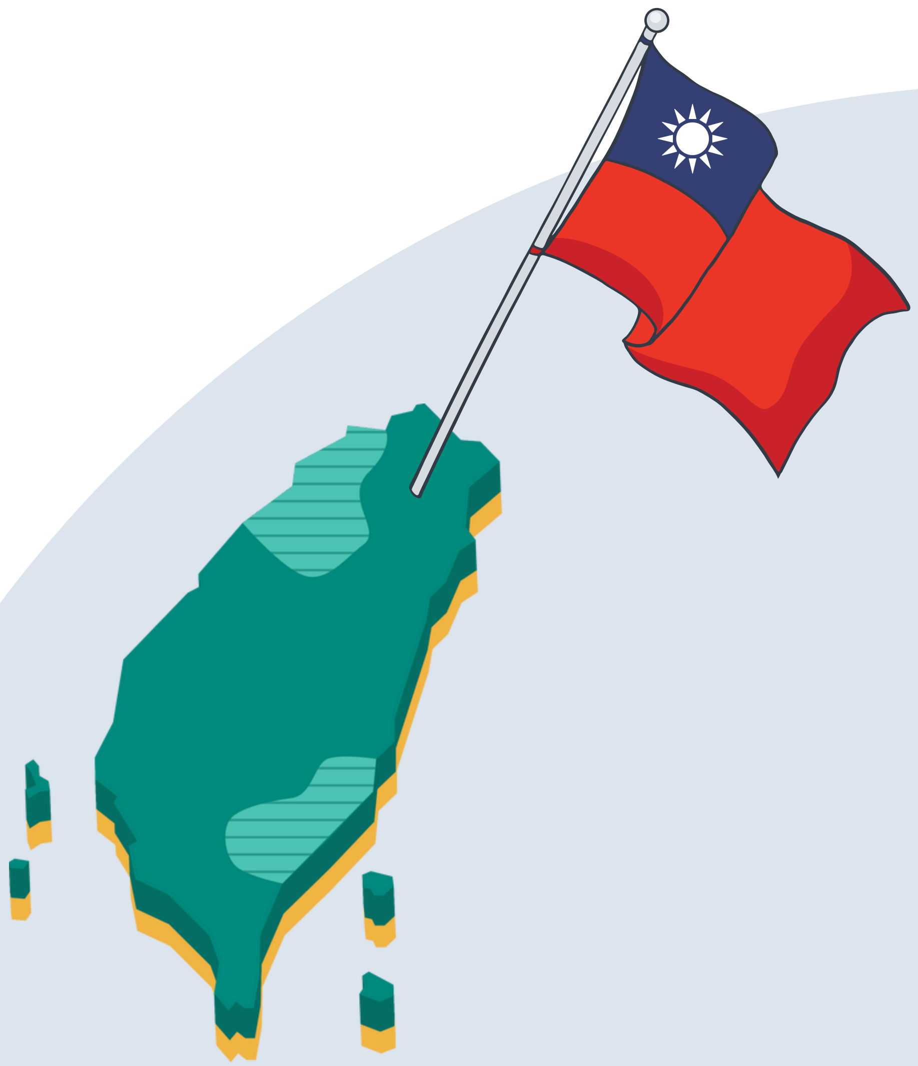
Ideology Prediction in Taiwanese Political News

A comparative study of machine learning, transformer-based models and knowledge distillation

June 6, 2025 | 邱秉辰、鄭博宇、劉士銓、林澐稚、林立潔



Introduction

Introduction

Media Bias in Taiwanese News

Related Work

- Two well-known models in political ideology classification:
  - Baly et al. (2020): fine-tunes BERT on 35k labeled articles from AllSides
  - Liu et al. (2022): **POLITICS**, pretrain RoBERTa on 2M+ news articles
- U.S. news is usually categorized by the **left–center–right** political spectrum.
  - Left: More liberal, supports social change, bigger government
  - Right: More conservative, favors tradition, smaller government
- In Taiwan, media ideology is shaped by **national identity**, not left–right politics.
  - Pro-green: Supports DPP, Taiwanese independence, opposes China’s policies.
  - Pro-blue: Supports KMT, opposes independence, tends to be pro-China.

Our goal: Ideology prediction in Taiwan

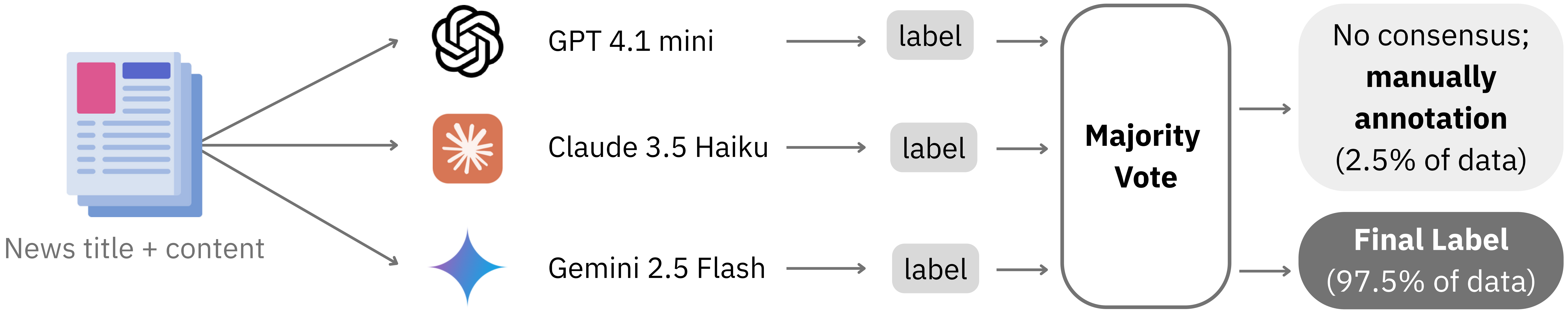
- We aims to build a Chinese BERT model to recognizes the ideology of Taiwanese political-news snippets
- Perform **knowledge distillation** to transfer ideological knowledge from **POLITICS**, enabling a Chinese BERT model to learn how to detect political bias in news content
- Our goal is to train a model that can understand the political stance of Taiwanese news — whether it's *pro-green*, *pro-blue*, or *neutral*

Dataset

Dataset

Data Collection and Annotation

- News data from major Taiwanese media outlets, including **TVBS**, **FTV**, **SET**, and **PTS** from **mid-April 2025** to **mid-May 2025**
  - Each article includes **the media source**, **news title**, **publication date**, and **full content**.
- Data annotation workflow



Data Overview

- We observe **noticeable differences in political bias** distribution across media outlets
  - **TVBS**: *pro-blue* (35%), *pro-green* (32%), and *neutral* (33%) articles
  - **SET** and **FTV**: have majority of *pro-green* articles, (54%) and (64%) respectively
  - **PTS**: has the highest proportion of *neutral* articles (56%)

Media	**Pro-Blue (%)**	**Pro-Green (%)**	**Neutral (%)**	**Total**
TVBS	388 (35%)	359 (32%)	363 (33%)	1110
SET	140 (15%)	504 (54%)	297 (32%)	941
PTS	37 (16%)	65 (28%)	132 (56%)	234
FTV	130 (14%)	574 (64%)	196 (22%)	900
Total	695 (22%)	1502 (47%)	988 (31%)	3185

Data Preprocessing

- We removed as much irrelevant information as possible from both the titles and content
  - **extraneous symbols**, **advertisements** at the end of articles, and **repeated mentions of media outlet names**
- Since some of our selected models operate on English data, we **translated** all articles from **Traditional Chinese** to **English** using GPT-4.1-mini
  - The model was asked maintain political tone, terminology, and journalistic style consistent with Taiwanese political news
- We also **removed articles with duplicate titles** or **identical content**
- The final dataset contained **3,166 unique** news articles

Method

Method

Task Definition

A **multi-class classification** problem

News title + content

Predict →

pro-blue, *pro-green*, or *neutral*

Method

Task Definition

A **multi-class classification** problem

News title + content

Predict

pro-blue, *pro-green*, or *neutral*

Formally,

$$x = (t, c)$$

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$$

Learn the function

$$f : X \rightarrow Y$$

$$Y = \{\textit{pro-blue},\ \textit{pro-green},\ \textit{neutral}\}$$

Baseline Models

Traditional ML Classifiers

- Logistic Regression
- Support Vector Machine

Transformer-based Models

Simplified Chinese

- BERT-base-Chinese (Google)
- RoBERTa-wwm-ext

Traditional Chinese

- BERT-base-Chinese (CKIP)

English (Translated)

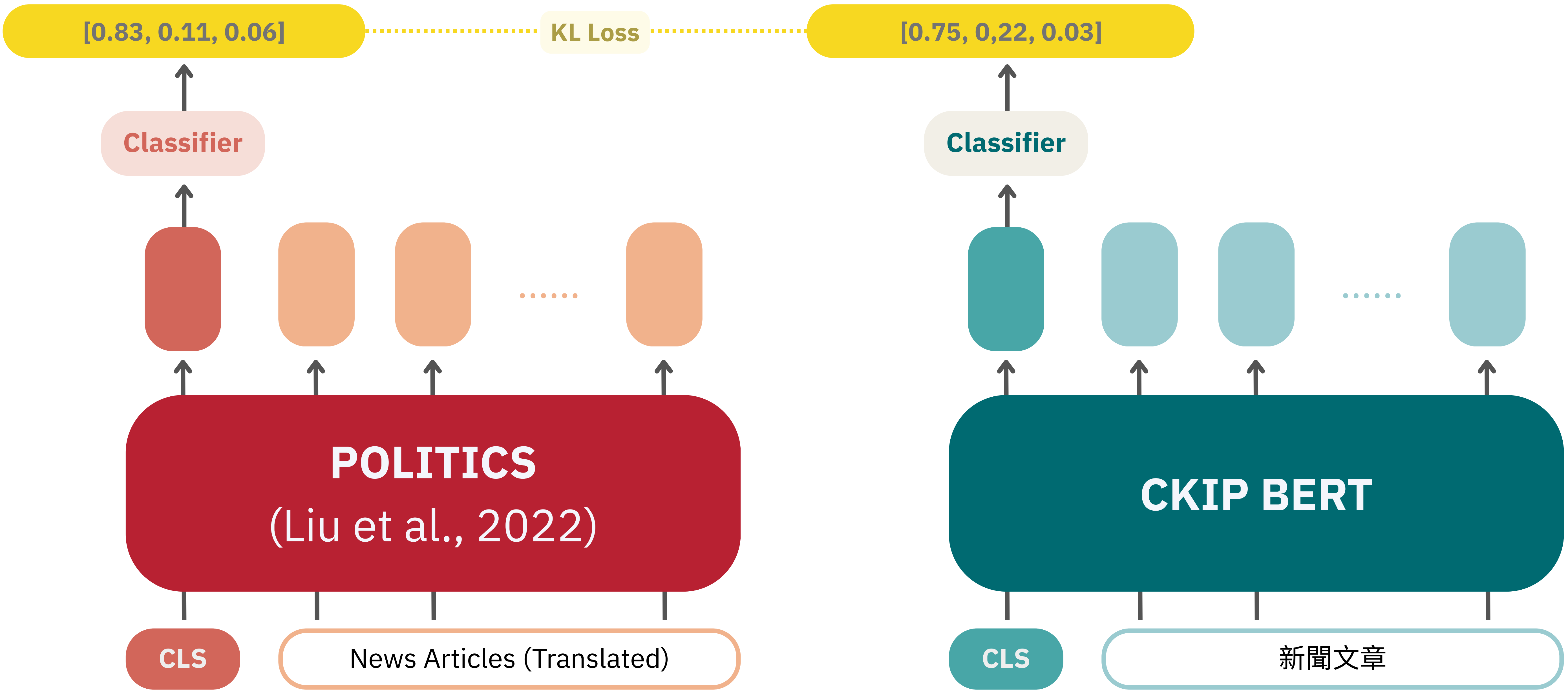
- BERT-base-uncased
- RoBERTa-base
- POLITICS

Proposed Models

- A **knowledge distillation** approach:
  - Let **POLITICS** be a **fixed teacher**
  - Let **CKIP BERT** be a **student**

- We propose the following two approaches:
  - Knowledge Distillation via **Soft Labels**
  - Knowledge Distillation via **CLS Embeddings**

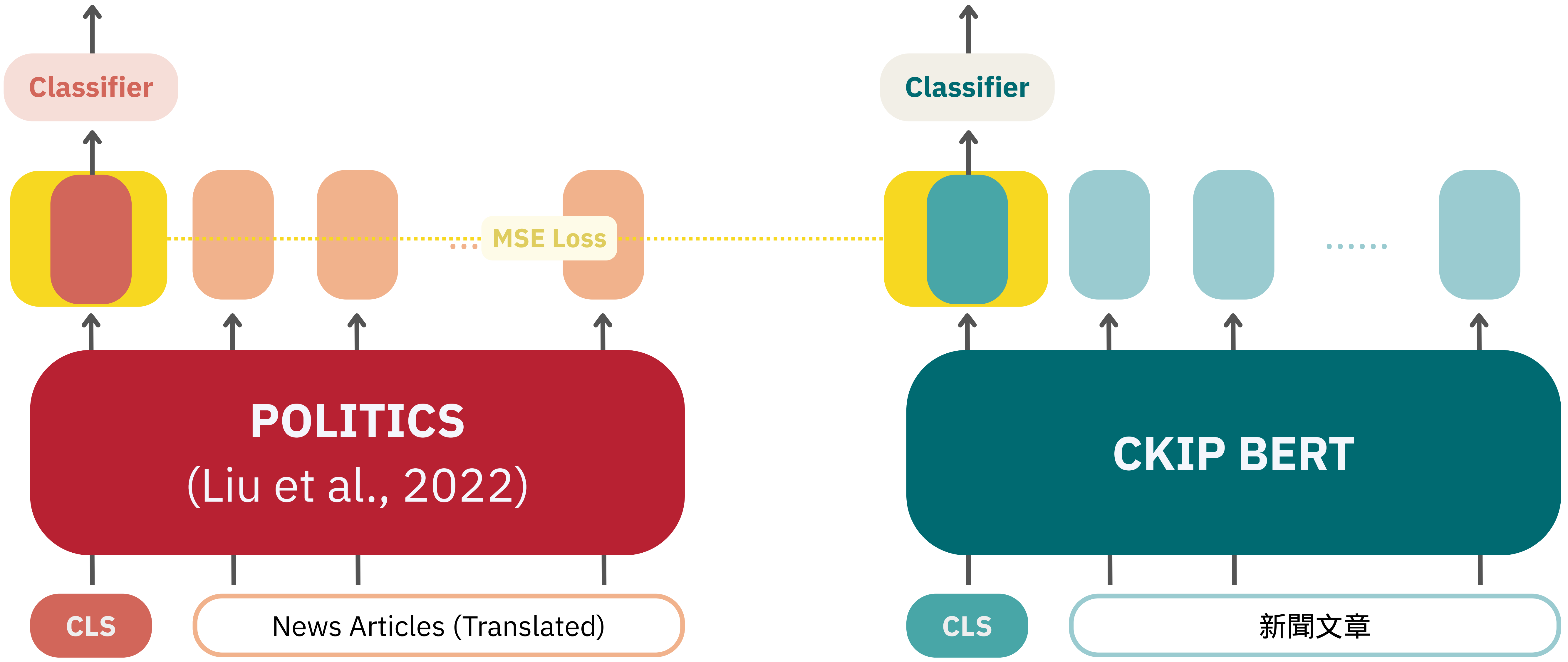
Knowledge Distillation via **Soft Labels**



The student is trained to mimic the teacher's soft-label predictions (Left / Center / Right) using KL divergence loss. We map these U.S. ideology labels to the Taiwanese categories either directly.

Method

Knowledge Distillation via **CLS Embeddings**



The student's [CLS] embedding (on the original Chinese text) is aligned with the teacher's [CLS] embedding (on the translated English text) using mean squared error (MSE) loss.

Loss Function of Knowledge Distillation

A **weighted loss function** with 2 components:

- Knowledge Distillation Loss
- Cross-Entropy loss

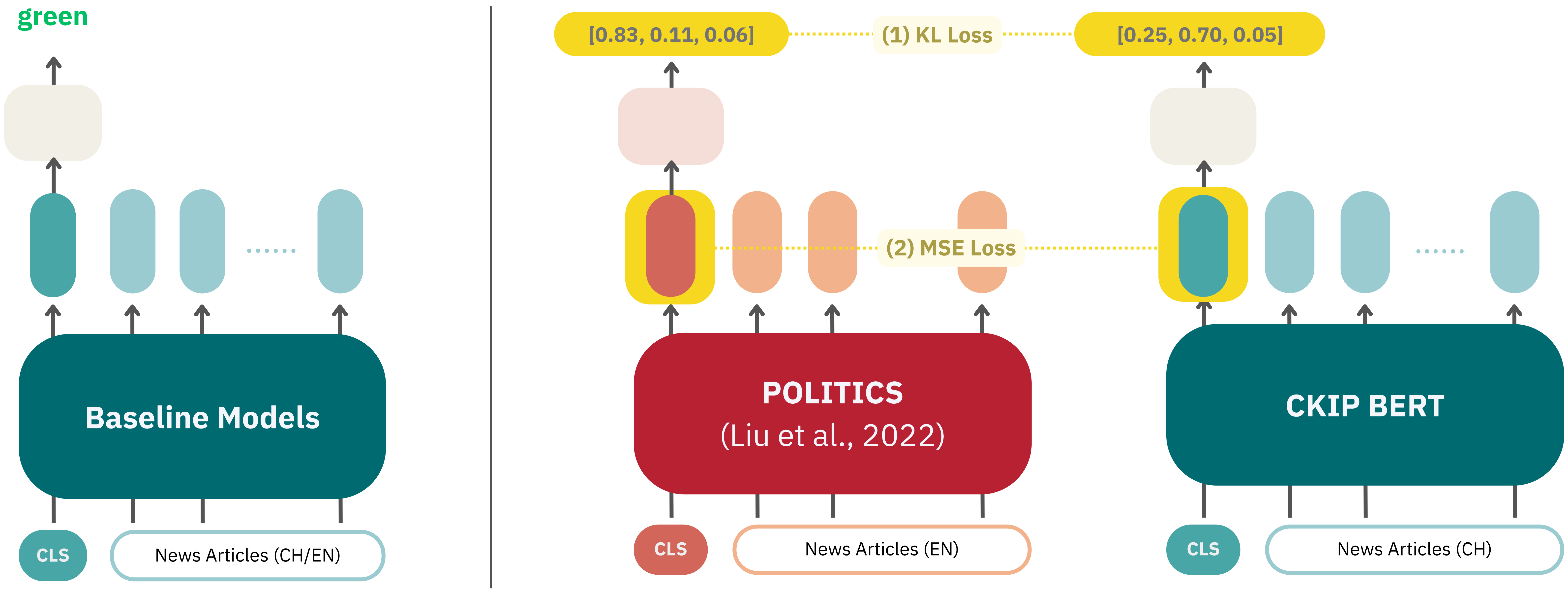
the ground-truth supervision

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{KD}} + (1-\alpha) \cdot \mathcal{L}_{\text{CE}}, \quad \alpha \in [0,1]$$

relative influence
of the teacher's guidance

Method

Comparison



Experiments & Results

Experiment Overview

- All models were assessed using **stratified 5-fold cross-validation.**
- Performance was reported as the **average across all folds**.
- Evaluation metrics (**macro-averaged**): **Accuracy, F1, Precision,** and **Recall**

Traditional Classifier Baselines

- **TF-IDF Vectorization** (min df = 5, max df = 0.8, stopword removal)
- Models:
  - Logistic Regression (liblinear, OvR)
  - SVM (RBF kernel, probability=True)
- **Class weights** used for data imbalance handling
- Implemented with Python 3 and scikit-learn library on Kaggle

Transformer-Based Baselines

- **Hyperparameters**

Hyperparameter	Value
Optimizer	AdamW
Learning Rate	0.00002
Batch Size	16
Epochs	5 (early stopping with patience 2)
Dropout / Warmup	0.1 / 0.1
Max Sequence Length	512

- All models used **class-weighted CE loss** to address imbalanced data
- Same set of hyperparameters in Knowledge Distillation setting
- Implemented with PyTorch & HuggingFace Transformers on Kaggle P100 GPU

Knowledge Distillation

- Model: **KD-CKIP (Soft / CLS)**
- All models used **class-weighted** CE loss to address imbalanced data
- Loss function during training:

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{KD}} + (1 - \alpha) \cdot \mathcal{L}_{\text{CE}}$$

  - We have $\alpha = 0.5$ in main experiment
  - Also conduct sensitivity test for $\alpha \in \{0, 0.2, 0.4, ..., 1\}$

Experiments & Results

Result #1: Overall Performance

Table 2: Performance comparison across models using title and content as input. All metrics are averaged over 5-fold cross-validation. Best results are **bolded**; second-best are underlined.

Category	Model	Macro Acc.	Macro-F1	Macro Precision	Macro Recall
Traditional ML	Logistic Regression	0.703	0.683	0.688	0.681
	SVM	0.705	0.687	0.691	0.688
Chinese	BERT	0.744	0.736	0.732	0.748
	RoBERTa	0.752	0.742	0.738	0.752
	CKIP BERT	0.767	0.755	0.755	<u>0.756</u>
English	BERT	0.738	0.725	0.725	0.727
	RoBERTa	<u>0.770</u>	<u>0.757</u>	<u>0.760</u>	0.755
	POLITICS	**0.776**	**0.761**	**0.770**	<u>0.756</u>
Ours (Distilled)	KD-CKIP (Soft)	0.751	0.743	0.738	0.754
	KD-CKIP (CLS)	0.755	0.748	0.744	**0.758**

KD-CKIP (CLS) achieves the **highest recall**, but their overall performance **falls short of several transformer-based baselines**.

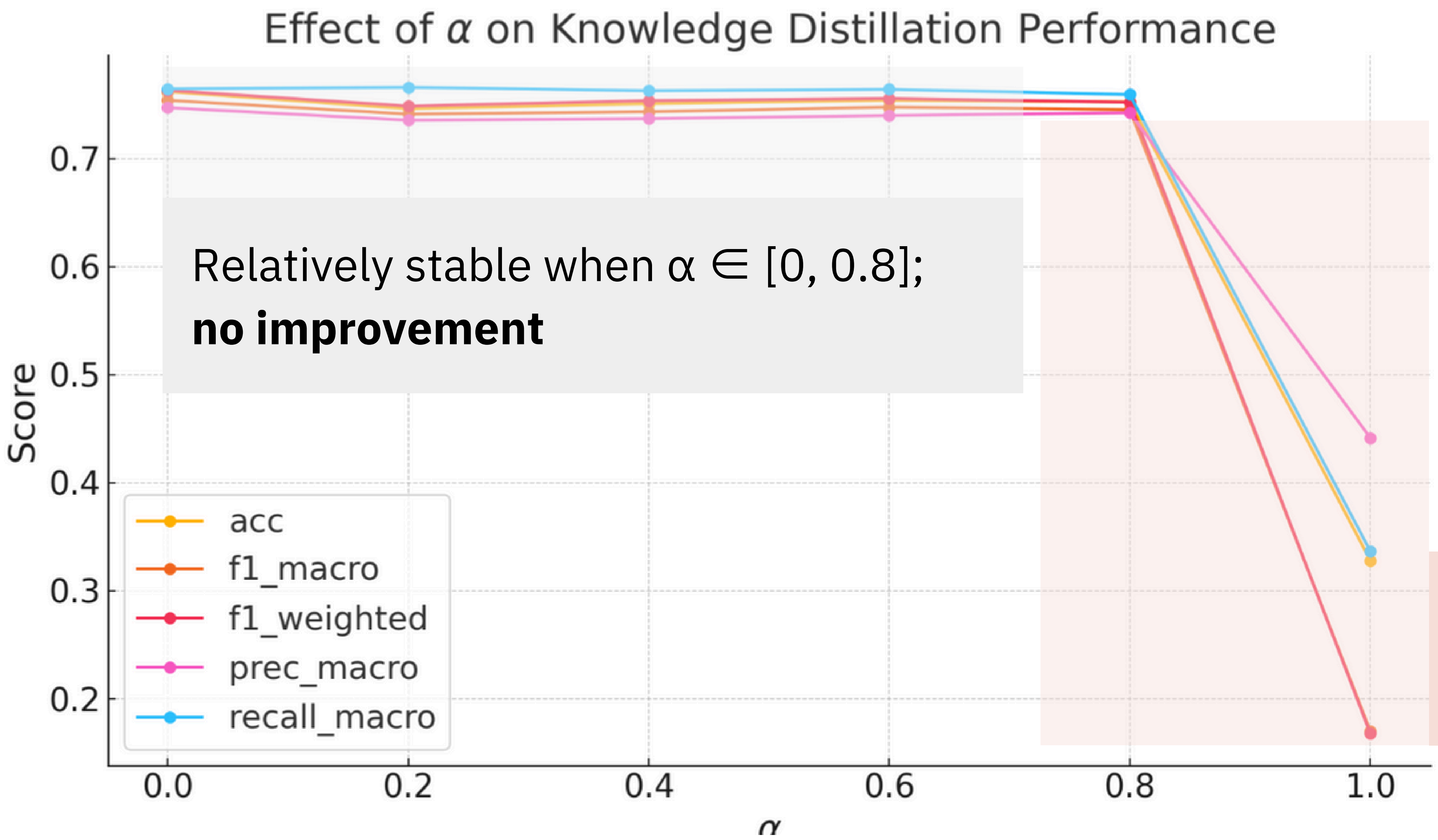
Transformer-based models **outperform traditional ML** baselines

CKIP BERT achieves the best performance among Chinese models

POLITICS achieves the overall best performance

Result #2: Effect of Distillation Weight on Loss Function

$$\mathcal{L}_{\text{total}} = \alpha \cdot \mathcal{L}_{\text{KD}} + (1 - \alpha) \cdot \mathcal{L}_{\text{CE}}$$



Relatively stable when α ∈ [0, 0.8];
no improvement

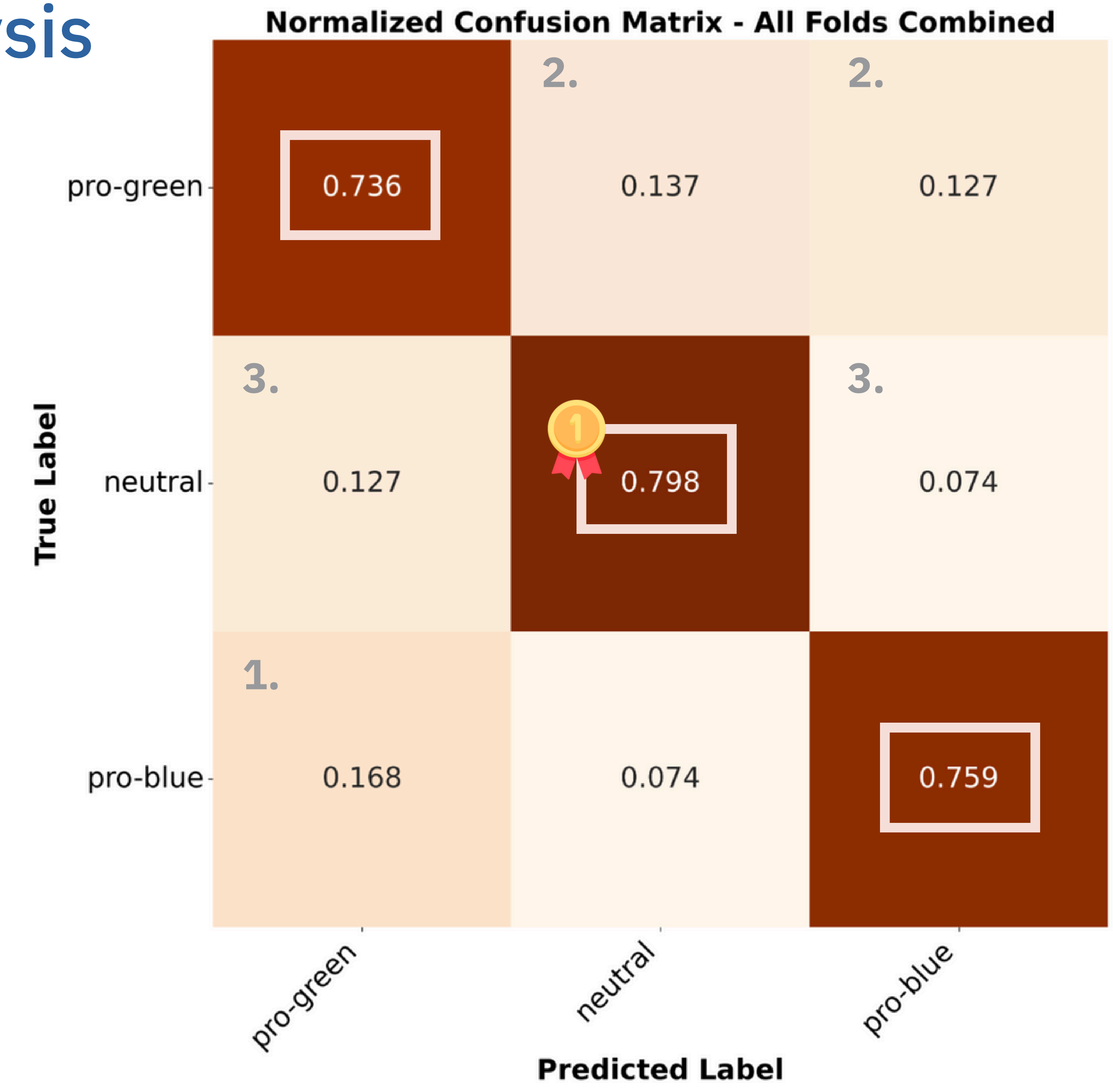
Performance **dropped sharply** when α = 1

We use standard 5-fold cross-validation instead of the stratified version in this experiment. Since our primary focus is on the effect of α, we consider this inconsistency with the main experiment acceptable.

Result #3: Confusion Matrix Analysis

Relatively balanced classification performance across the three ideological categories, with the **highest accuracy** observed in the **neutral class**.

1. *Pro-blue* articles are most frequently misclassified as *pro-green* **(16.8%)**
2. *Pro-green* instances are confused with both *neutral* (13.7%) and *pro-blue* (12.7%)
3. *Neutral* is more often misclassified as *pro-green* **(12.7%)** than *pro-blue* (7.4%)



Case Study: World Master Game

總統府突派蕭美琴出席世壯運
藍議員爆「維安大亂」：見黃仁勳就蹭

True: pro-blue　True: pro-blue　TVBS

2025世界壯年運動會明天將於台北大巨蛋盛大開幕，由全球AI龍頭輝達（NVIDIA）創辦人暨執行長黃仁勳親自擔綱象徵「科技」領域的聖火手，總統賴清德不出席開幕典禮，不過副總統蕭美琴將代為出席。據了解，北市府2月25日就發函給總統府，不過府在昨天才回覆總統不克前往，導致隨扈警力大亂，遭國民議員抨擊，「見名人就蹭、見掌聲就衝」的急功近利。

據了解，北市府是在2月25日世壯運正式發函邀請正副總統出席，5月15日賴清德報告國政繁忙不出席，5月16日8點宣布黃仁勳擔任聖火手，總統府遂於上午9點多通知蕭美琴出席，導致從北市警力的部署、國安維安作業、軍方人力動員，三個方面都要重新調整，蕭美琴隨扈，被召回；維安人力臨時被迫動起來都得銷假。

賴清德缺席世壯運遭轟　游淑慧：不是要讓世界看見台灣?

True: pro-blue　Predicted: pro-green　TVBS

2025世界壯年運動會今（17）日將在台北大巨蛋開幕，總統賴清德因公不克出席，遭藍營痛批，台灣好不容易爭取到的國際運動盛事，賴身為總統卻不願盡義務，看不起世壯運、看不起2萬多名選手。不過，民進黨發言人吳崢對此稱，這是在幫台北市長蔣萬安轉移近期的政治壓力，轉移世壯運亂象，喊話不需要有這些政治攻防；國民黨北議員游淑慧則說，過去民進黨一直說要讓國際看見台灣，世壯運有107個國家的選手來到台灣，賴總統卻不出席？不是國民黨把這件事政治化，而是總統把它變成一個笑話。

針對賴清德不出席世壯運，遭藍營痛批，吳崢16日在TVBS政論節目《少康戰情室》中表示，不用把這件事太政治化解讀，世壯運這場合類似世大運，總統如果可以來露臉沒有任何不妥或扣分，並透露，最後副總統蕭美琴決定參加，經他詢問總統府，確實之前台北市教育局有寄一份公文給總統府秘書長，邀請總統、副總統參加，府方表示總統這次是另有公務行程，因為總統、副總統行程較忙，大概會協調到很接近活動時間的時候，再看是誰參加，這其實也是美事一樁，副總統過去在國際上曝光度或知名度也是很高。

Case Study: World Master Game

總統該去世壯運? 趙怡翔批政治操作：30年僅1國家首長參與

True: pro-green

Predicted: pro-green

TVBS

雙北世壯運於今天在台北大巨蛋盛大開幕，台北市政府昨公布壓軸聖火代表，由輝達（NVIDIA）創辦人暨執行長黃仁勳擔綱「科技」聖火手，隨後副總統蕭美琴宣布出席開幕。引發藍營質疑總統府「蹭」黃仁勳聲量。民進黨台北市議員趙怡翔則指出，過去30年僅有一位國家首長親自出席世壯運開幕，批評藍營此舉為政治操作，並強調應聚焦賽事本身。

民進黨台北市議員趙怡翔昨晚於臉書發文，表示部分藍營人士批評賴清德因公務未能出席開幕是明顯的政治操作。他指出，過去30年世壯運僅有一位國家首長參與開幕，其餘皆由地方首長代表，並舉出1998年波特蘭、2002年墨爾本、2005年愛德蒙頓、2009年雪梨及2013年杜林等歷屆案例說明，都一致由市長或州長開場。

Conclusion

Summary of our findings

- **Transformer-based models** significantly **outperform traditional classifiers**
  - POLITICS model achieving the best performance on English-translated data.
  - Among Chinese models, CKIP BERT performs strongly
- **Knowledge distillation** strategies performed competitively but did not outperform direct model fine-tuning
- **Confusion matrix analysis** revealed **asymmetries in misclassification**, suggesting certain ideological distinctions remain challenging
  - between *pro-blue (true)* and *pro-green (predict)*
  - between *neutral (true)* and *pro-green (predict)*

Limitations and Future Work

Dataset was mostly annotated by large language models

- **Domain experts are still needed to capture the nuanced pattern**

Data imbalance and limited generalizability

- **Broaden the scope to include a wider array of media outlets and political events**

Knowledge distillation had unexpected performance

- **Consider multilingual pre-trained models**

Thanks for Listening!

邱秉辰、鄭博宇、劉士銓、林澐稚、林立潔