

---

---

# Final Presentation

— Jiliang (Eric) Li, Kangbai (Jack) Yan, Youngbin  
(Kyle) Kwon, Liam Betts, Kun Peng —

---

---

# Problem Statement

## Concept Drift

- New Topics / Keywords
- New Summary / Writing Styles
- New Meanings to Words

The **New York Times**  
in the past



The **New York Times**  
in the present



It's helpful if a model can be updated, via **Incremental Learning**

- First fine-tuned on a large quantity of news summaries
- Then incrementally updated on small batches of up-to-date news

# Problem Statement

## Concept Drift

- New Topics / Keywords
- New Summary / Writing Styles
- New Meanings to words

The **New York Times**  
in the past



The **New York Times**  
in the present



## Incremental Learning

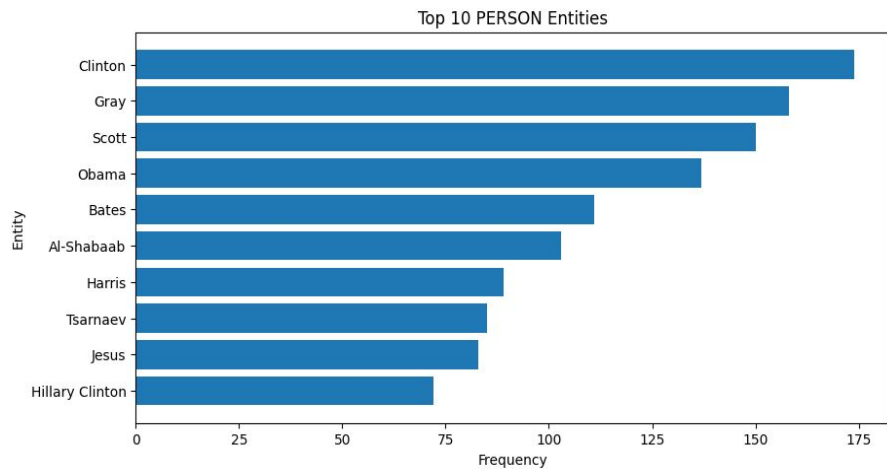
CNN/DailyMail Dataset

- First fine-tuned on a large quantity of news summaries
- Then incrementally updated on small batches of up-to-date news

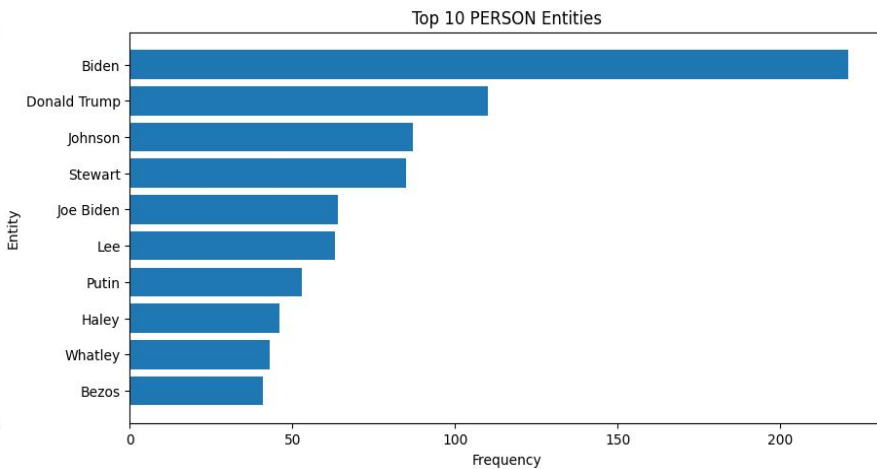
NewsAPI

# Motivation - Domain Shift Visualized via NER

CNN/DailyMail Dataset

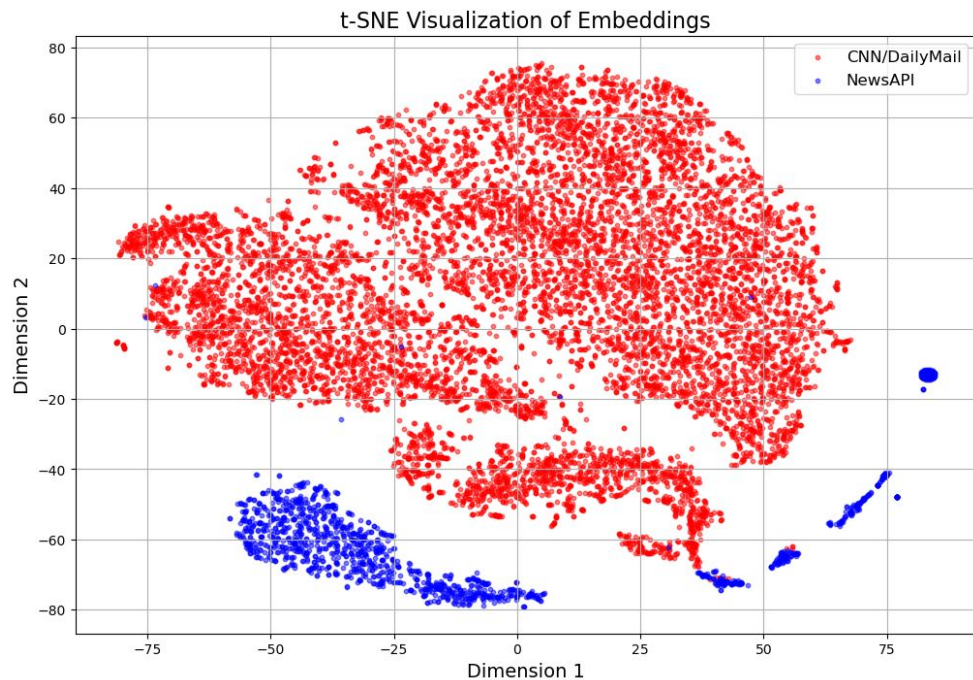


NewsAPI Data (~500 articles)



# Motivation - Domain Shift Visualized via Embedding

- The **last\_hidden\_layer** of a **bart-large-cnn** as the embedding of news articles.
- T-SNE clustering



# Data Exploration - CNN / DailyMail Dataset

	article	highlights	id
0	LONDON, England (Reuters) -- Harry Potter star...	Harry Potter star Daniel Radcliffe gets £20M f...	42c027e4ff9730fbb3de84c1af0d2c506e41c3e4
1	Editor's note: In our Behind the Scenes series...	Mentally ill inmates in Miami are housed on th...	ee8871b15c50d0db17b0179a6d2beab35065f1e9
2	MINNEAPOLIS, Minnesota (CNN) -- Drivers who we...	NEW: "I thought I was going to die," driver sa...	06352019a19ae31e527f37f7571c6dd7f0c5da37
3	WASHINGTON (CNN) -- Doctors removed five small...	Five small polyps found during procedure; "non...	24521a2abb2e1f5e34e6824e0f9e56904a2b0e88
4	(CNN) -- The National Football League has ind...	NEW: NFL chief, Atlanta Falcons owner critical...	7fe70cc8b12fab2d0a258fababf7d9c6b5e1262a

## Dataset Split

## Number of Instances in Split

Train

287,113

Validation

13,368

Test

11,490

## Feature

## Mean Token Count

Article

781

Highlights

56

Total Articles: 287113

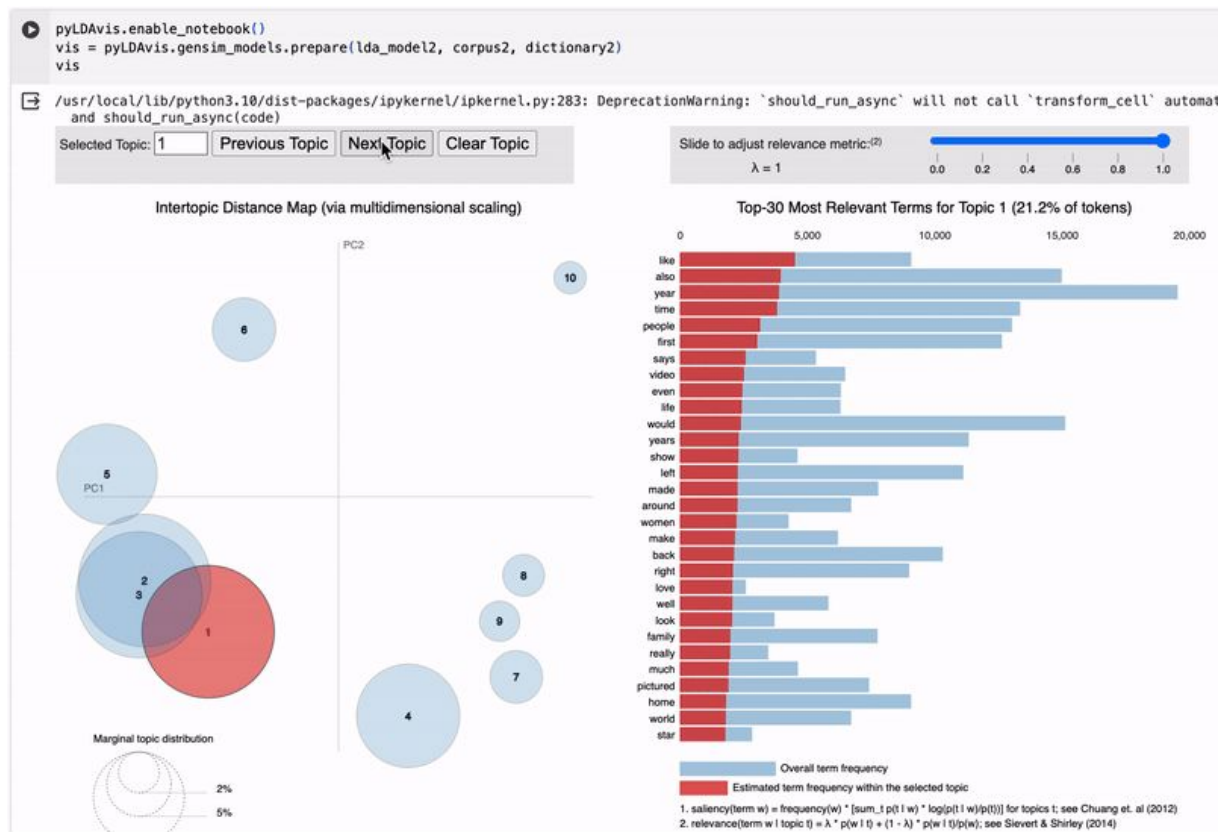
Average Article Length: 691.8703263175126 words

Average Summary Length: 51.574101486174435 words



# Topic Modeling

- Added “said” to stopwords
- Used Latent Dirichlet Allocation (LDA) model
- Typical topics fairly evident:  
government/politics,  
economy/business,  
entertainment,  
public health, and  
sports.





# NER (Named Entity Recognition) using spaCy

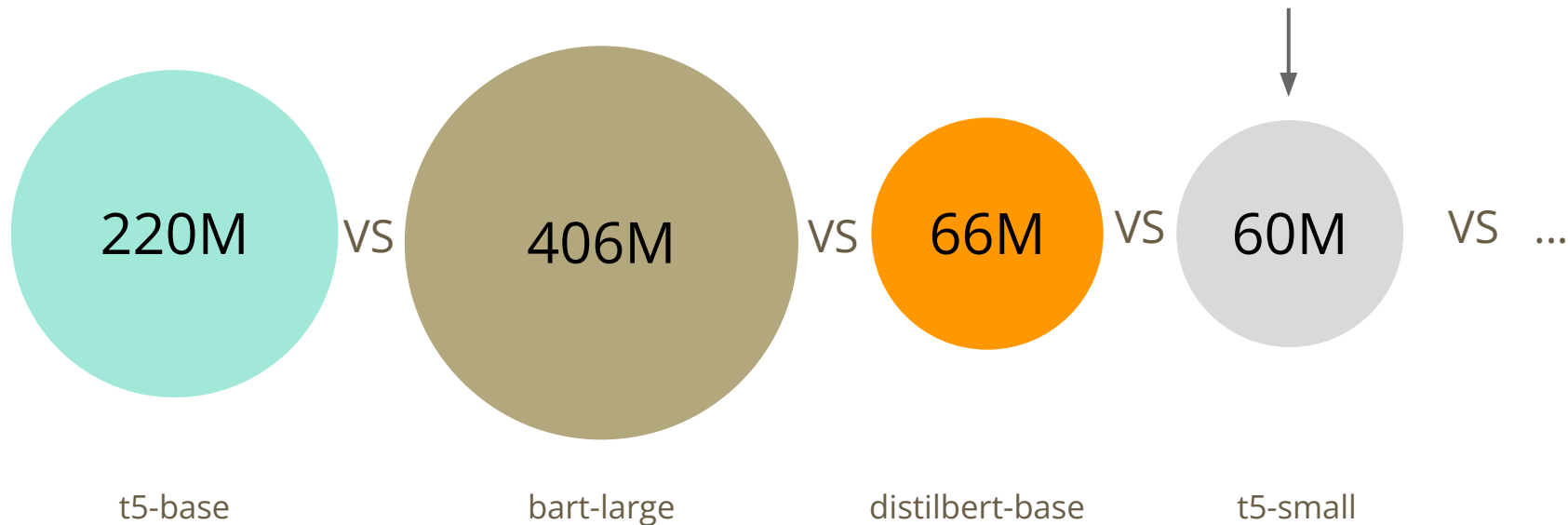
- **PERSON:** People, including fictional.
- **NORP:** Nationalities, religious or political groups.
- **FAC:** Facilities, like buildings, airports, highways, bridges, etc.
- **ORG:** Organizations, including companies, agencies, institutions, etc.
- **GPE:** Geopolitical entities, like countries, cities, states.
- **LOC:** Non-GPE locations, mountain ranges, bodies of water.
- **PRODUCT:** Objects, vehicles, foods, etc. (Not services.)
- **EVENT:** Named hurricanes, battles, wars, sports events, etc.

nose at the top," Clinton PERSON said, vowing to be a "champion" of "everyday Ame  
conomic themes. She has taken on a range of issues that most appeal to liberals, such as t  
edia to defend the 2010 DATE Dodd-Frank Wall Street ORG reform law, which coi  
Bill Daley PERSON, a longtime Clinton PERSON ally and former Obama GPE  
crisis, Daley PERSON said, "that's a problem." He continued: "My guess is she'll have e  
to Massachusetts GPE Democratic NORP Sen. Elizabeth Warren PERSON a  
ressive views. Former Republican NORP New Hampshire GPE Sen. Judd Gregg  
ange stems from the fact that "the center of the Democratic Party ORG has moved ver  
aid. "I presume she's going to get the nomination, but she may be contested from the left."  
ig banks as they tour the early presidential states. Former Maryland GPE Gov. Marti  
dependent who caucuses with Democrats NORP, are all testing the waters. They have  
reet and toughen financial regulations. None of these potential candidates will be able to co  
o embrace progressive policies. After Clinton PERSON's announcement Sunday D  
Hillary Clinton PERSON and other candidates laying out their platforms and hearing whet

# Model Selection

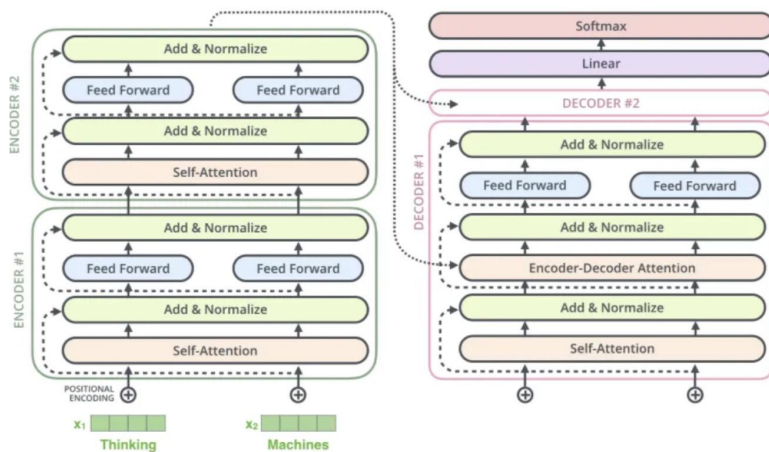
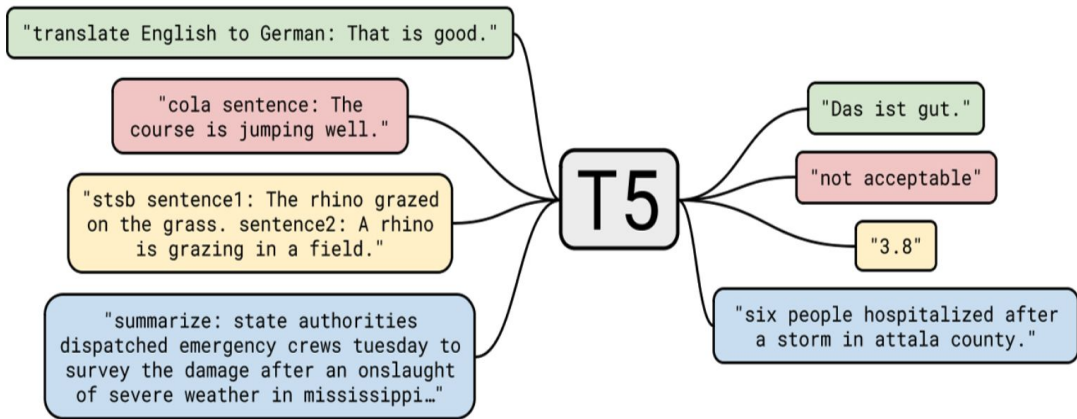


A balance between performance (seq2seq) and memory efficiency (size)



Num of Parameters

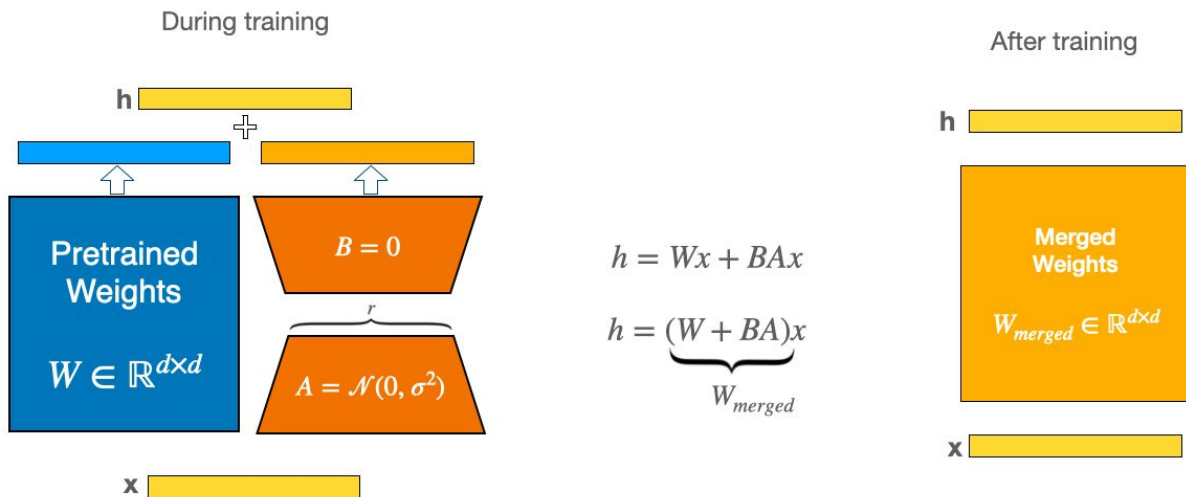
# Model Selection



A diagram of our text-to-text framework. Every task we consider—including translation, question answering, and classification—is cast as feeding our model text as input and training it to generate some target text. This allows us to use the same model, loss function, hyperparameters, etc. across our diverse set of tasks. It also provides a standard testbed for the methods included in our empirical survey. “T5” refers to our model, which we dub the “Text-to-Text Transfer Transformer”.

# LoRA for Model Fine-Tuning

- Reduce the number of trainable parameters
- Low inference latency
- Model updates are decoupled from the pre-trained weights



# Metrics

## Recall-Oriented Understudy for Gisting Evaluation (ROGUE)

- Compares an automatically produced summary or translation against reference (high-quality and human-produced) summaries or translations.
- Different variants:
  - **ROUGE-N**: Measures the overlap of n-grams between the system and reference summaries.
  - **ROUGE-L**: Based on the Longest Common Subsequence (LCS), it accounts for sentence-level structure similarity.

# Metrics

$$F1 = \frac{2 \times Precision \times Recall}{(Precision + Recall)}$$

**Precision:** # n-grams in S that appear also in A / # n-grams in S

**Recall:** # n-grams in A that appear also in S / # n-grams in A

**Article (A):** The cat is on the mat.

**Summary (S):** The cat and the dog.

**ROUGE-1:** Precision = 3/5 = **0.6**

Recall = 3/6 = **0.5**

F1-score = **0.55**

**ROUGE-2:** Precision = 1/4 = **0.25**

Recall = 1/5 = **0.20**

F1-score = **0.22**

**ROUGE-L:** Precision = 3/5 = **0.6**

Recall = 3/6 = **0.5**

F1-score = **0.55**

# News Retrieval Pipeline

- Fully automated and hosted on **Google Cloud**.
- Scheduled to automatically query **NewsAPI**, a service that gathers the latest news from numerous sources (NYT, WSJ, etc).

The logo for News API, featuring the word "News" in white sans-serif font on a blue background, followed by "API" in blue sans-serif font inside a white square on the same blue background.

News API

# News Retrieval Pipeline

- Processes the latest news articles, on average numbering **~500 articles / day**.
  - NewsAPI provides a “description” for each article, which is used as the summary
  - The article url is scraped for its content.
  - Then, the summary and content is grouped as data and fed into our model for the incremental learning to take into effect.



```
← Top business headlines in the US right now →
GET https://newsapi.org/v2/top-headlines?country=us&category=business&apiKey=API_KEY

{
  status: "ok",
  totalResults: 67,
  - articles: [
    - {
      - source: {
        id: "cnn",
        name: "CNN"
      },
      author: "Ramishah Maruf, Pete Muntean, Gregory Wallace",
      title: "Alaska Airlines CEO says company found loose bolts on 'many' Boeing Max 9s - CNN",
      description: "Alaska Airlines CEO Ben Minicucci revealed the carrier found 'some loose bolts on many' Boeing 737 Max 9s in an interview for 'NBC Nightly News with Lester Holt' scheduled to air Tuesday.",
      url: https://www.cnn.com/2024/01/23/business/alaska-airlines-ceo-boeing-max-9-loose-bolts/index.html,
      urlToImage: https://media.cnn.com/api/v1/images/stellar/prod/240123161554-ben-minicucci-ceo-alaska-airlines-09152022.jpg?c=16x9&q=w_800,c_fill,
      publishedAt: "2024-01-24T00:22:30Z",
      content: "Alaska Airlines CEO Ben Minicucci revealed the carrier found some loose bolts on many Boeing 737 Max 9s in an interview for NBC Nightly News with Lester Holt scheduled to air Tuesday.\r\nIts the CEOs fi... [+3538 chars]"
    },
    - {
      - source: {
        id: null,
        name: "Yahoo Entertainment"
      },
      ..
    }
  ]
}
```



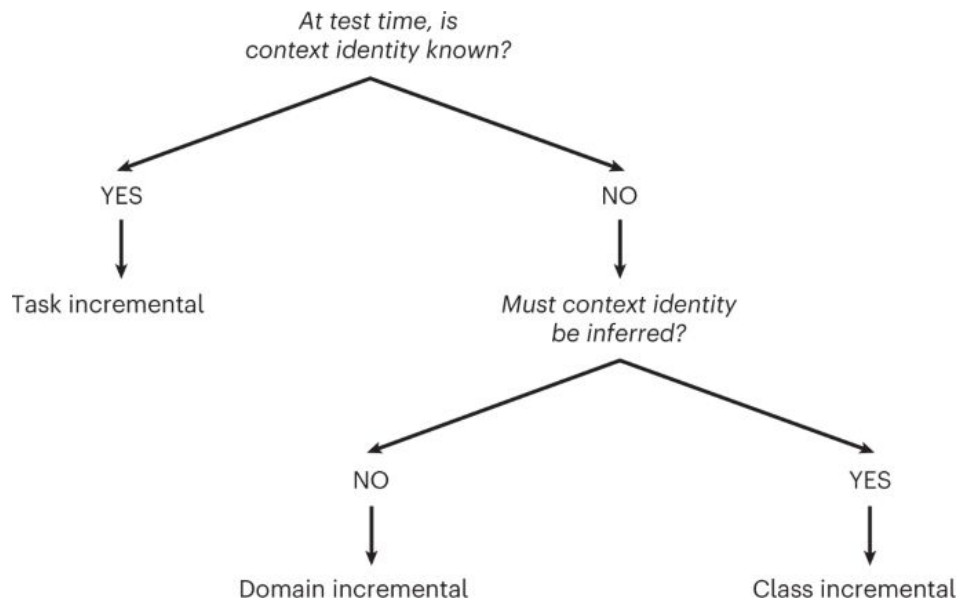
# News Retrieval Pipeline

- The news sources are carefully selected and filtered
  - Considering the different biases and reputation
- The pipeline also has built-in error handling
  - Able to continue processing data even if faulty links or data is encountered.

# Literature Review on Incremental Learning

Three types of Incremental Learning:

1. Task Incremental
2. **Domain Incremental (Our task)**
3. Class Incremental



# But, IL/CL Comes with Catastrophic Forgetting

## Catastrophic Forgetting

- Models might perform worse on older datasets as they are more and more fine-tuned on new datasets.

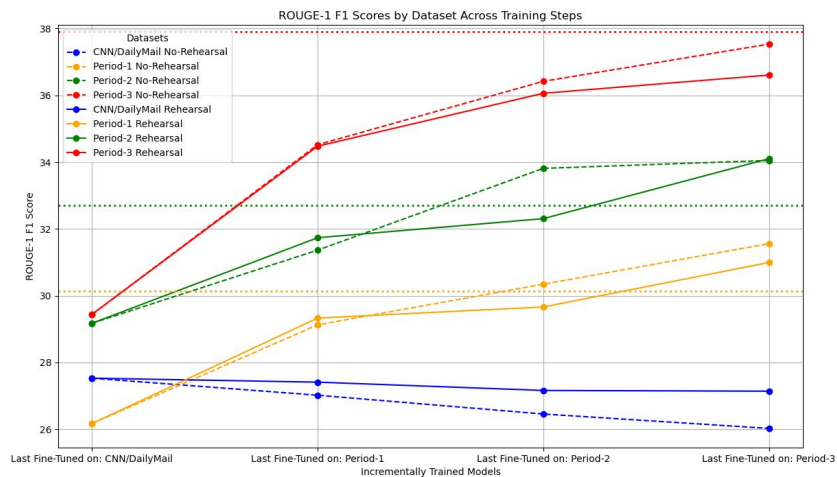
## Solutions Attempted

- **Experimental Rehearsal / Replay:** storing previous samples and reusing them when training on new samples.
- **Elastic Weight Consolidation (EWC):** While learning a new task, EWC protects the performance on an old task by constraining the parameters to stay in a region of low error for the old task.

# Experiments with Rehearsal/Replay

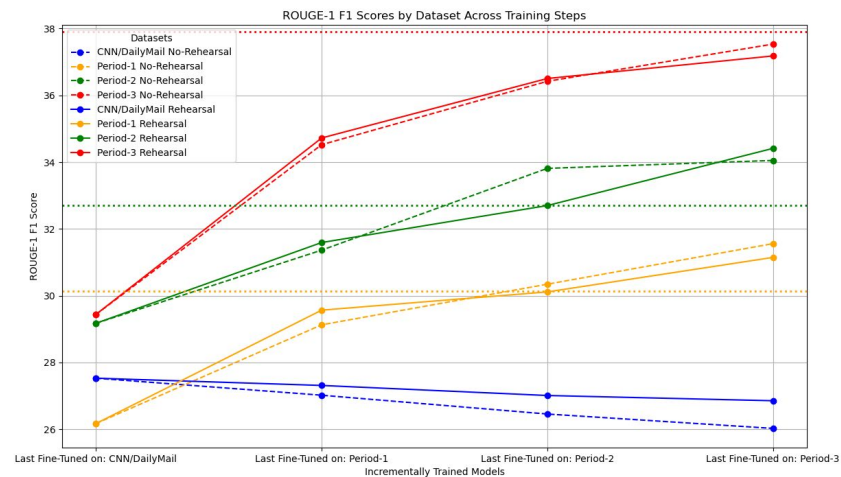
Rehearsal vs. No Rehearsal

(33% old data)



Rehearsal vs. No Rehearsal

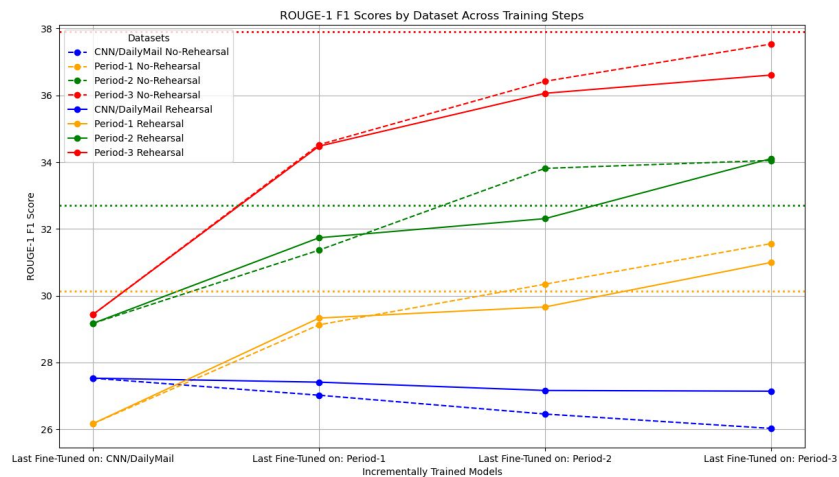
(16% old data)



# Experiments with Rehearsal/Replay

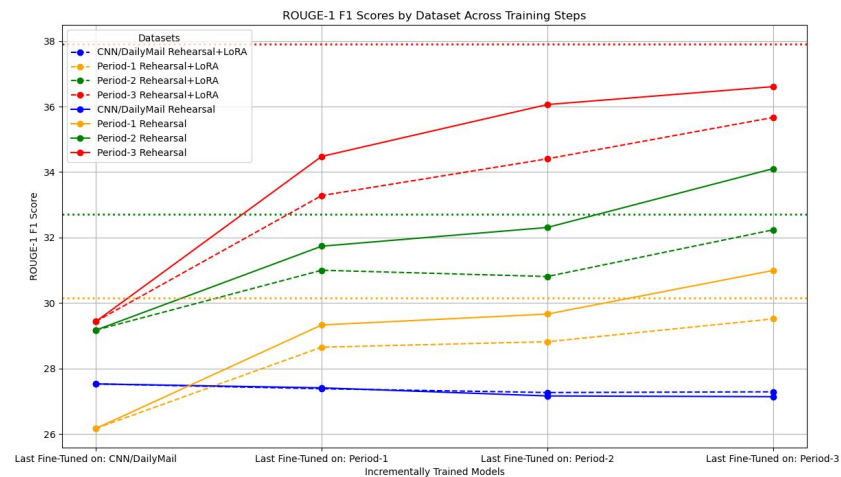
Rehearsal vs. No Rehearsal

(33% old data)



Rehearsal vs. LoRA Rehearsal

1764 mb vs. 938 mb

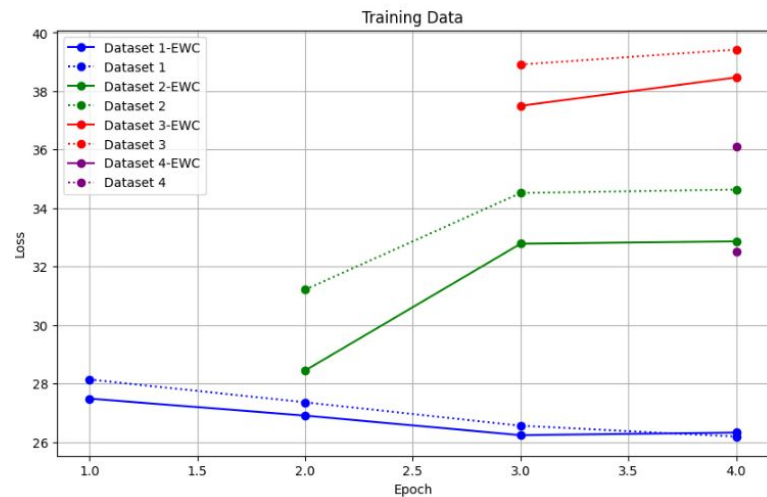


# Experiments with EWC

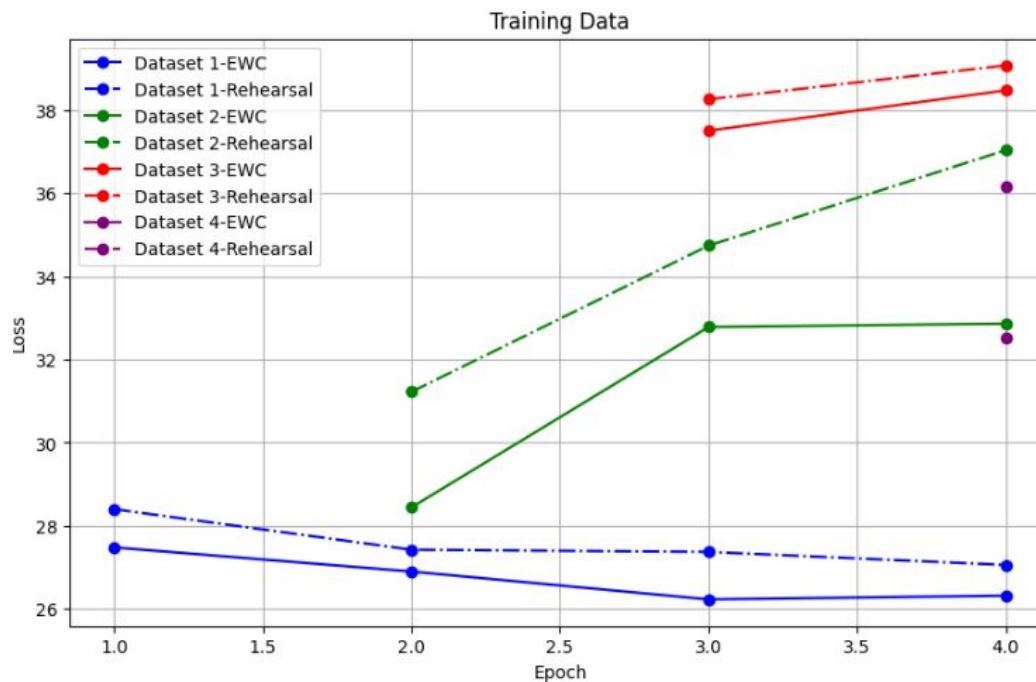
## EWC Quadratic Penalty

$$\mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \sum_i \frac{\lambda}{2} F_i(\theta_i - \theta_{A,i}^*)^2 \quad (3)$$

where  $\mathcal{L}_B(\theta)$  is the loss for task B only,  $\lambda$  sets how important the old task is compared to the new one and  $i$  labels each parameter.



# Experiments with EWC



# Rehearsal vs EWC

- Regularization (EWC) doesn't perform as well as Rehearsal method, as we expected before the experiment based on related works.
- Adopt the **REHEARSAL** strategy for continual learning for the final deployed model



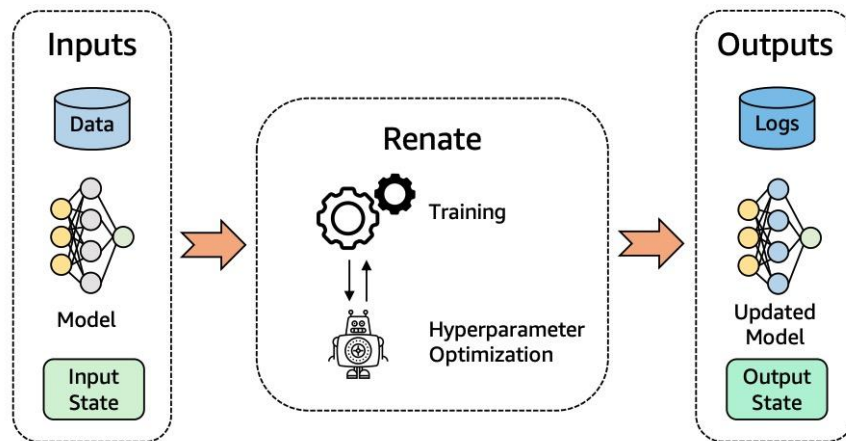
# Continual Learning API

## Avalanche: an End-to-End Library for Continual Learning

Powered by ContinualAI



powered by



# Trial on Avalanche and Renate

- Find compatible CL packages for Huggingface and PyTorch
- Run some demos of CL at Colab

```
from typing import Callable, Dict, Optional

import torch
from torchmetrics import Accuracy
from torchvision.transforms import transforms

from renate import defaults
from renate.benchmark.datasets.vision_datasets import TorchVisionDataModule
from renate.benchmark.models.mlp import MultiLayerPerceptron
from renate.benchmark.scenarios import ClassIncrementalScenario, Scenario
from renate.models import RenateModule

def data_module_fn(data_path: str, chunk_id: int, seed: int = defaults.SEED) -> Scenario:
    """Returns a class-incremental scenario instance.

    The transformations passed to prepare the input data are required to convert the data to
    PyTorch tensors.
    """
    data_module = TorchVisionDataModule(
        data_path,
        dataset_name="MNIST",
        val_size=0.1,
        seed=seed,
    )

    class_incremental_scenario = ClassIncrementalScenario(
        data_module=data_module,
        groupings=((0, 1, 2, 3, 4), (5, 6, 7, 8, 9)),
        chunk_id=chunk_id,
    )
    return class_incremental_scenario

def model_fn(model_state_url: Optional[str] = None) -> RenateModule:
    """Returns a model instance """
```

```
Avg acc: 37.74333333333333
Training on task: 1
Train Epoch: 1 Loss: 1.987624
Train Epoch: 2 Loss: 1.593181
Testing on task: 0
Test set: Average loss: 0.0062, Accuracy: 5666/10000 (57%)

Testing on task: 1
Test set: Average loss: 0.0034, Accuracy: 7701/10000 (77%)

Testing on task: 2
Test set: Average loss: 0.0114, Accuracy: 1164/10000 (12%)

Avg acc: 48.43666666666667
Training on task: 2
Train Epoch: 1 Loss: 2.034442
Train Epoch: 2 Loss: 1.997389
Testing on task: 0
Test set: Average loss: 0.0082, Accuracy: 5225/10000 (52%)

Testing on task: 1
Test set: Average loss: 0.0068, Accuracy: 4019/10000 (40%)

Testing on task: 2
Test set: Average loss: 0.0042, Accuracy: 6989/10000 (70%)

Avg acc: 54.10999999999999
```

# Trial on Avalanche and Renate

An unfortunate result:

- Conflicts with the version of multiple python packages
- Did not meet our expectations (improvement, workload)
- Few references online (haven't updated since Jan 2023)

# HuggingFace Space

Online and running:

<https://huggingface.co/spaces/liamvbetts/bart-news-summary-v1>

```
app.py 2 x requirements.txt
app.py > ...
1 import gradio as gr
2 |
3 from transformers import AutoTokenizer, AutoModelForSeq2SeqLM
4
5 tokenizer = AutoTokenizer.from_pretrained("liamvbetts/bart-large-cnn-v4")
6 model = AutoModelForSeq2SeqLM.from_pretrained("liamvbetts/bart-large-cnn-v4")
7
8 def summarize(article):
9     inputs = tokenizer(article, return_tensors="pt").input_ids
10    outputs = model.generate(inputs, max_new_tokens=128, do_sample=False)
11    summary = tokenizer.decode(outputs[0], skip_special_tokens=True)
12    return summary
13
14 # Create Gradio interface
15 input_text = gr.Textbox(lines=10, label="Input Text")
16 output_text = gr.Textbox(label="Summary")
17
18 gr.Interface(
19     fn=summarize,
20     inputs=input_text,
21     outputs=output_text,
22     title="News Summary App",
23     description="Enter a news text and get its summary."
24 ).launch()
```

Spaces | liamvbetts / bart-news-summary-v1

like 0 Running

## News Summary App

Enter a news text and get its summary.

Input Text

Clear Submit

Summary

# HuggingFace Space

## News Summary App

Enter a news text, search for news articles, or load a random article.

Search for News

taylor swift

Model Name

liamvbetts/bart-news-summary-v1

Search News Article

Summarize

Label

**Elon Musk Says Taylor Swift Should Be Concerned About Jet Tracking**

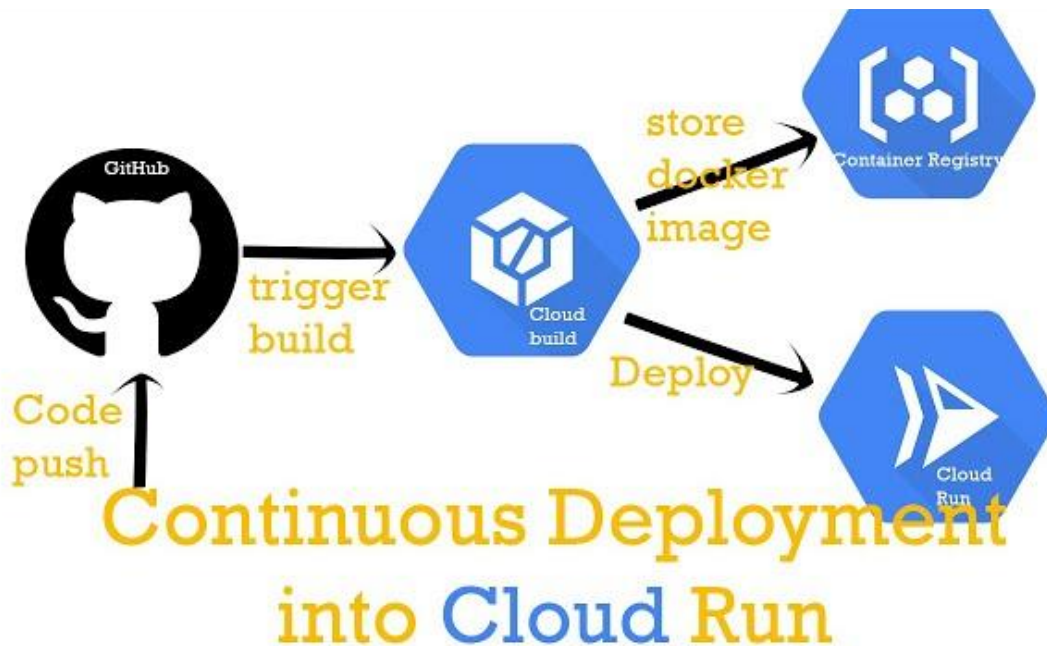
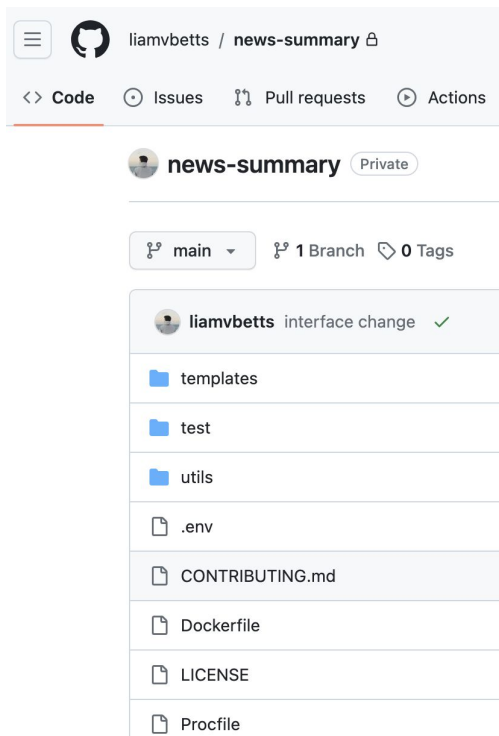
Summary

Elon Musk called Jack Sweeney "an awful human being" for tracking celebrities' private planes . Swift's attorney sent Sweeney a cease-and-desist letter in December, calling for him to stop "stalking and harassing behavior" Sweeney says he never meant any harm to the people whose jets he tracked, and he created the accounts because he

Input Text

Elon Musk has entered the chat. On Tuesday, the billionaire spoke out about Taylor Swift threatening legal action against a college student named Jack Sweeney for posting the locations of celebrities' private planes online. "Sweeney is an awful human being," Musk wrote on X in response to a post made by a journalist, Ian Miles Cheong, about the news. "Taylor Swift is right to be concerned." Sweeney is an awful human being. Taylor Swift is right to be concerned. Sweeney told Business Insider over text that Musk's post was typical of the billionaire. "Why does he get involved in literally everything?" Sweeney said. The college student brought up his history with Musk in a response to Musk on X. "Let's remember you offered me 5k to stop," Sweeney wrote. "Also threatened to sue me, and here we are a year later." Musk's comment came after news broke of Swift's attorney sending Sweeney a cease-

# Cloud Run with Continuous Deployment



# Flask App

- Updated / organized UI
- Button to fetch today's fine-tuned model
- Now calculates ROUGE scores between NewsAPI highlights and generated summaries

## News Summary App

spring break

liamvbetts/bart-news-summary-v1

**Article Title:** How To Find The Best Travel Insurance For Your Spring 2024 Trip

Spring break 2024 might be dangerous. Here's how to protect yourself. The State Department poured a bucket of cold water on spring break 2024 when it recently issued warnings for several popular spring destinations. It also got everyone to start thinking about travel insurance,

**State Department issued travel warnings for spring break 2024 . Travel is risky, particularly during the 2024 spring break season . You should be thinking of protecting yourself . Travel insurance is one way to do that . Here's how to protect yourself this year and in the years to come**

**Results:**

Word Count - Article: 512

Word Count - Summary: 52

## News Summary App

Enter topic, or press below for a random article.

**Article Title:** Alex Pereira explains waving off referee Herb Dean, meme celebration after Jamahal Hill knockout - MMA Fighting

get in the way of another highlight reel. . . . . retained his light heavyweight title with a brutal first-round knockout of Jamahal Hill in the main event of UFC 300 on Saturday at T-Mobile Arena in Las Vegas. Adding another knockout to his résumé is nothing new for Pereira, but the manner in which this one came about

**Article Highlights:**

Alex Pereira had some curious theatrics before and after his stunning knockout of Jamahal Hill at UFC 300.

**Model Summary:**

Alex Pereira retained his light heavyweight title with a brutal first-round knockout of Jamahal Hill . Pereira was seen waving off referee Herb Dean as the official looked to step in to address a low kick from Hill . Pereira was seen

**ROUGE1:** 0.171 0.389 0.237

**ROUGE2:** 0.100 0.235 0.140

**ROUGE L:** 0.171 0.389 0.237







# A Showcase of IL/CL's Effectiveness

Rouge scoring averaged over ~500 news articles released in 2024

	T5 + Incremental Learning	T5
Rouge-1	0.151605	0.098356
Rouge-2	0.110561	0.077415
Rouge-l	0.142509	0.093274

# A Case Study on IL/CL's Effectiveness

**Article:** Five crew members onboard the submersible Titan were probably killed instantly in a “catastrophic implosion” as it descended to the wreck of the Titanic two miles below the surface of the Atlantic ocean, US Coast Guard officials announced on Thursday. A large debris field containing multiple sections of the vessel was spotted earlier in the day by a remotely operated vehicle (ROV) scouring the seabed near the Titanic wreck site 400 miles south of St John’s, Newfoundland, officials said at an afternoon press conference in Boston.

**T5 + IR:** US Coast Guard officials announced Thursday five crew members onboard the submersible Titan were probably killed instantly in a “catastrophic implosion” as it descended to the wreck of the Titanic two miles below the surface of the Atlantic Ocean . A large debris field containing multiple sections of the vessel was spotted earlier in the day by a remotely operated vehicle (ROV) (0.686046511627907, 0.6588235294117647, 0.6046511627906976)

**T5:** US Coast Guard officials announced on Thursday that five crew members onboard the submersible Titan were killed instantly in a “catastrophic implosion” as it descended to the wreck of the Titanic two miles below the surface of the Atlantic ocean. Five crew members onboard the submersible Titan were probably killed instantly in a “catastrophic implosion” as it descended to the wreckage, US Coast Guard officials said. Five crew members onboard the submersible Titan were probably killed instantly in a float. (0.5232558139534884, 0.4235294117647059, 0.5116279069767442)

①

②

③

# Division of Labor

- **Eric:** Experiments with **rehearsal**, **LoRA** & hyperparameter fine-tuning, domain shift exploration & visualization
- **Jack:** Experiments with rehearsal & **EWC**, **literature review** on IL/CL, Huggingface framework integration
- **Kyle:** NewsAPI integration, automated **data pipeline** implementation, data preprocessing & filtering
- **Liam:** Data exploration & visualization, model fine tuning, HuggingFace deployment, **Flask app**, automated data pipeline implementation
- **Kun:** Literature Review on IL/CL, model selection, **Avalanche & Renate** implementation

**Thank you!**