

Reminders

- ▶ HW2 due 2/18 (postponed due to delays in publishing the assignment)
- ▶ PRA3 due 2/22
- ▶ Course project progress report I due 2/27
- ▶ Come to OH for course project discussion!

Artificial Intelligence Methods for Social Good

Lecture 9

Case Study: NewsPanda: AI for Conservation-Related Media Monitoring

17-537 (9-unit) and 17-737 (12-unit)

Instructor: Fei Fang

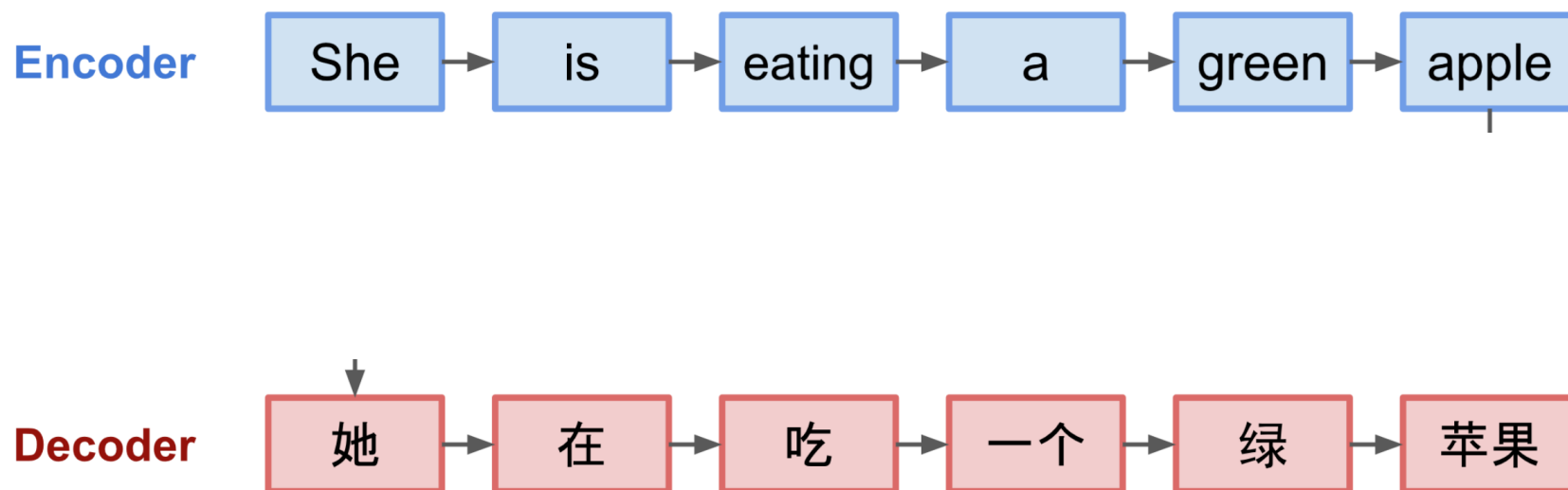
feifang@cmu.edu

Outline

- ▶ Transformer revisited
- ▶ NewsPanda
- ▶ NewsSerow
- ▶ Discussion

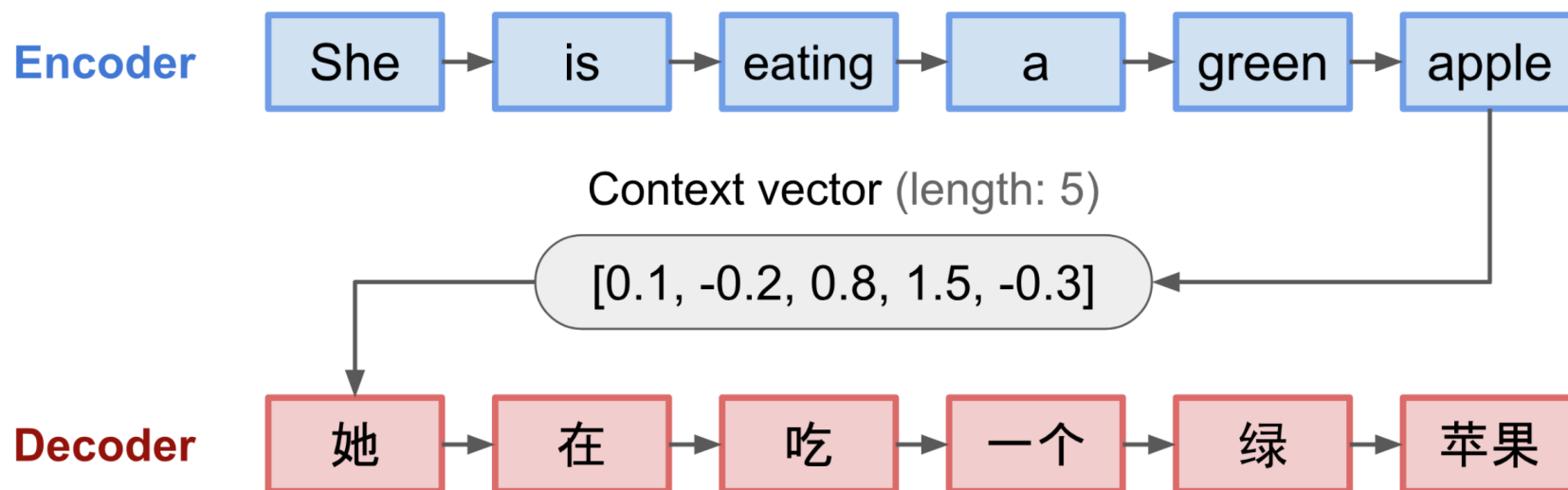
Recap: Attention (in the context of translation)

- ▶ Each output token depends on input tokens differently
- ▶ Intuitively, calculate the importance weight for each of the source token for current predicting token



Recap: Attention (in the context of translation)

- ▶ Source sequence: $x = [x_1, x_2, \dots, x_m]$
- ▶ Target sequence: $y = [y_1, y_2, \dots, y_m]$
- ▶ h_i : encoder state at the i th position in source
- ▶ $s_t = f(s_{t-1}, y_{t-1}, c_t)$: decoder hidden state for the output word at position t where c_t is context vector



Recap: Attention (in the context of translation)

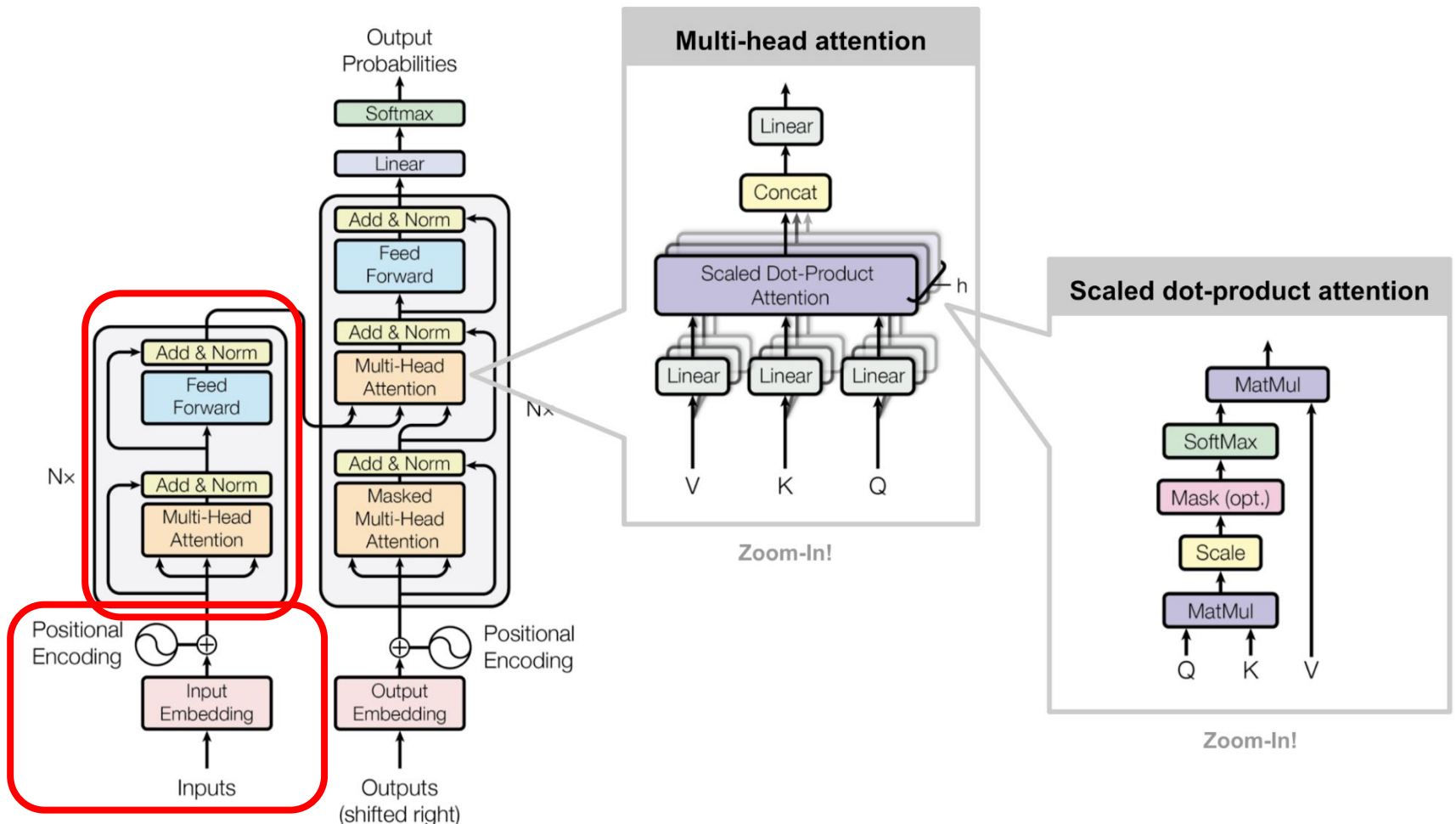
- ▶ Source sequence: $x = [x_1, x_2, \dots, x_m]$
- ▶ Target sequence: $y = [y_1, y_2, \dots, y_m]$
- ▶ h_i : encoder state at the i th position in source
- ▶ $s_t = f(s_{t-1}, y_{t-1}, c_t)$: decoder hidden state for the output word at position t where c_t is context vector

$$c_t = \sum_i \alpha_{t,i} h_i$$

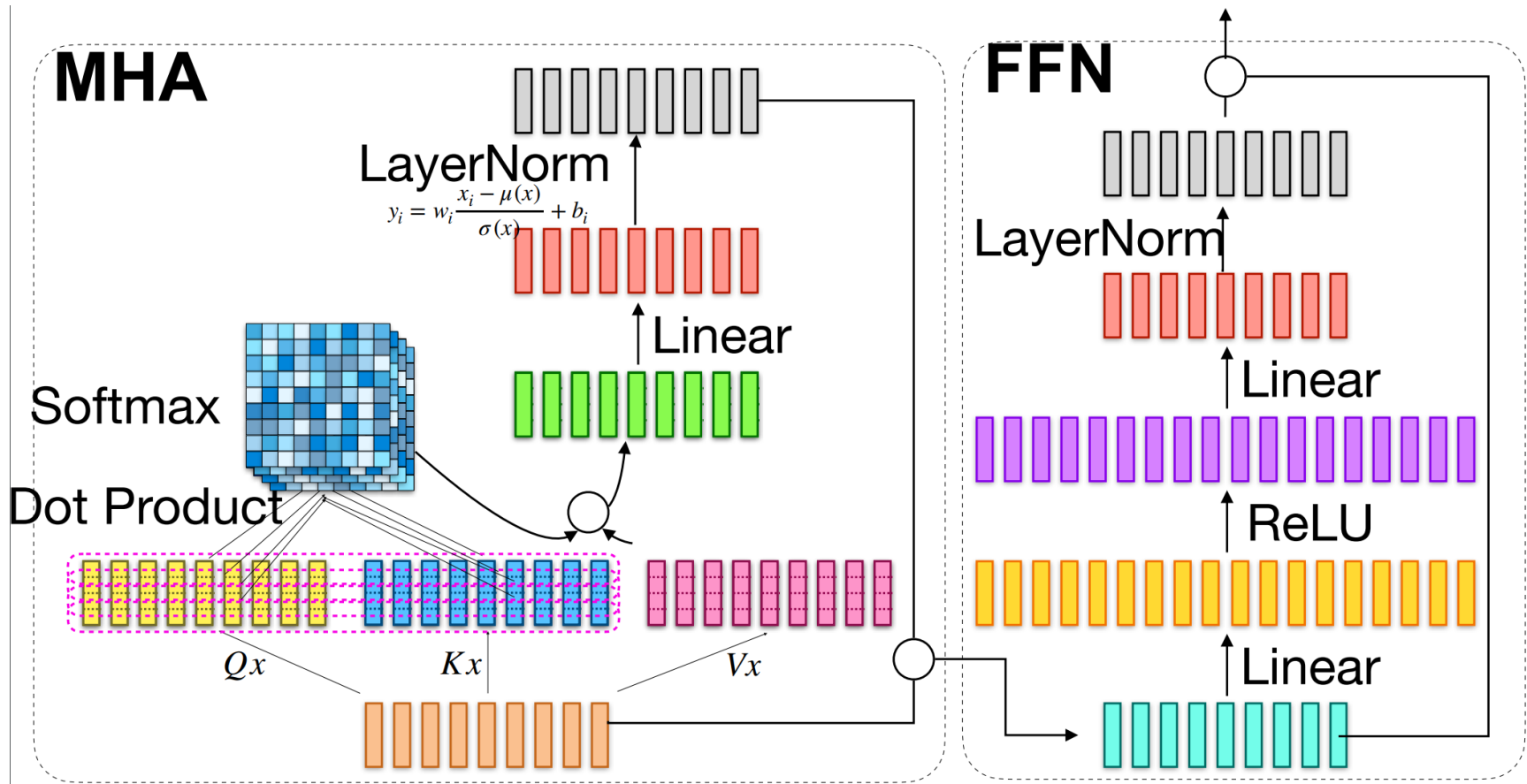
$$\text{where } \alpha_{t,i} = \frac{\exp(\text{score}(s_{t-1}, h_i))}{\sum_{i'} \exp(\text{score}(s_{t-1}, h_{i'}))}$$

$$\text{Scaled Dot-Product Attention: } \text{score}(s_t, h_i) = \frac{s_t^T h_i}{\sqrt{d}}$$

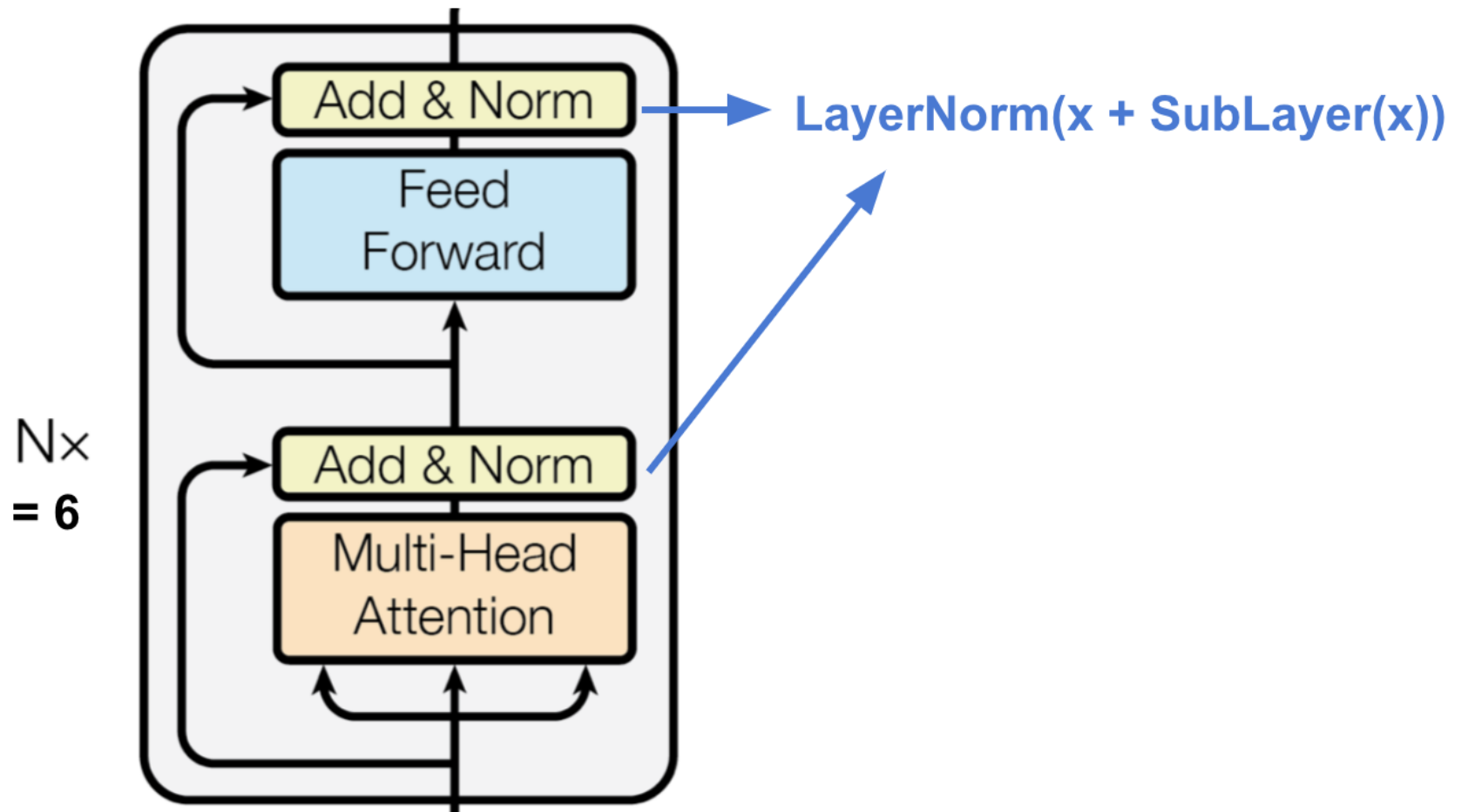
Recap: Transformer Architecture Full



Transformer Encoder

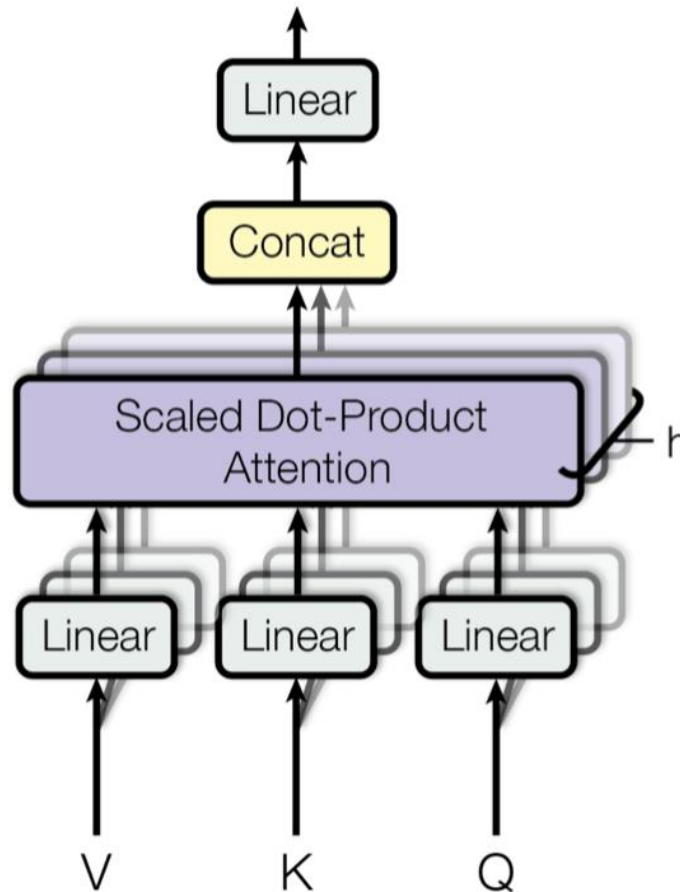


Transformer Encoder

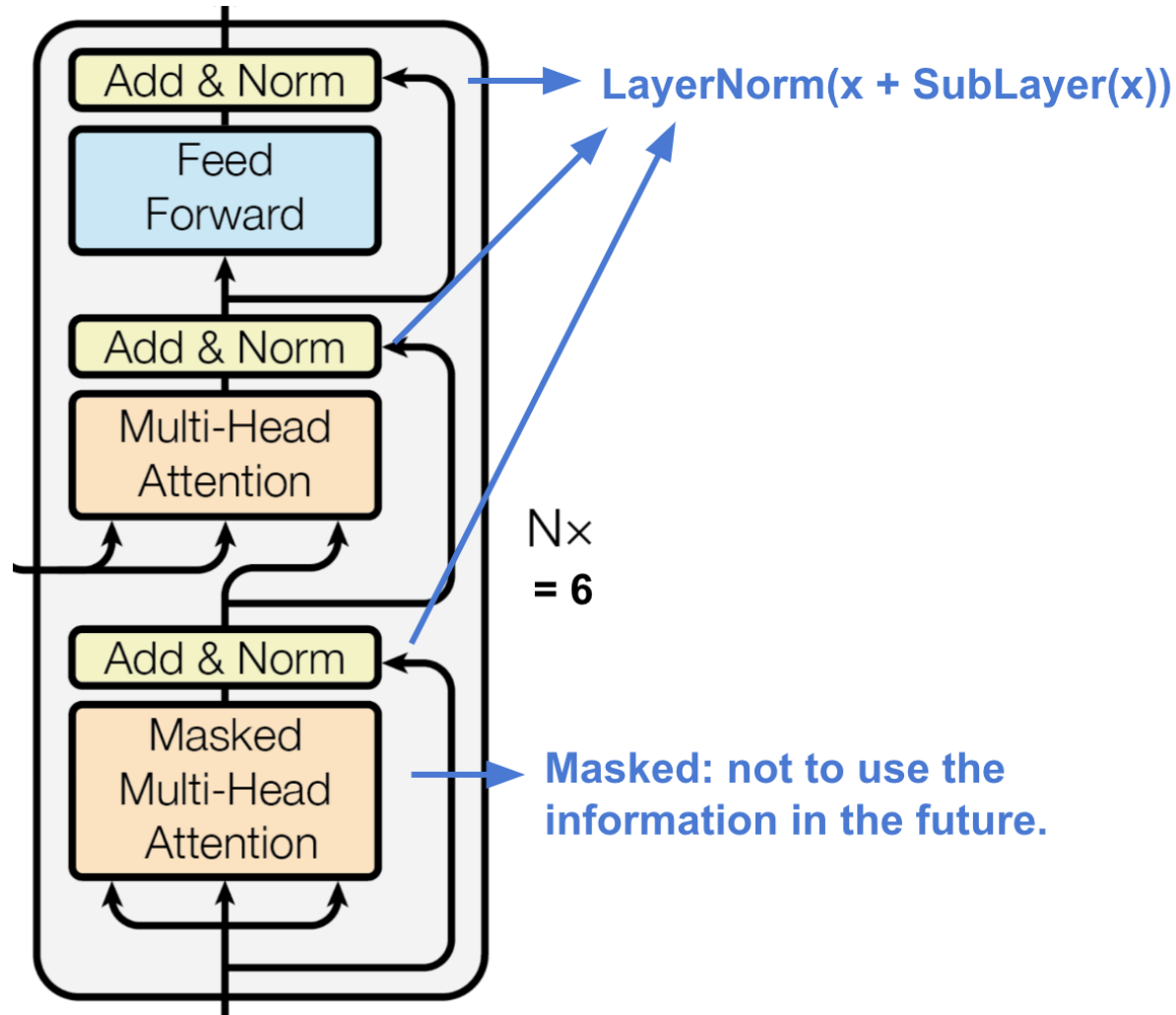


Transformer

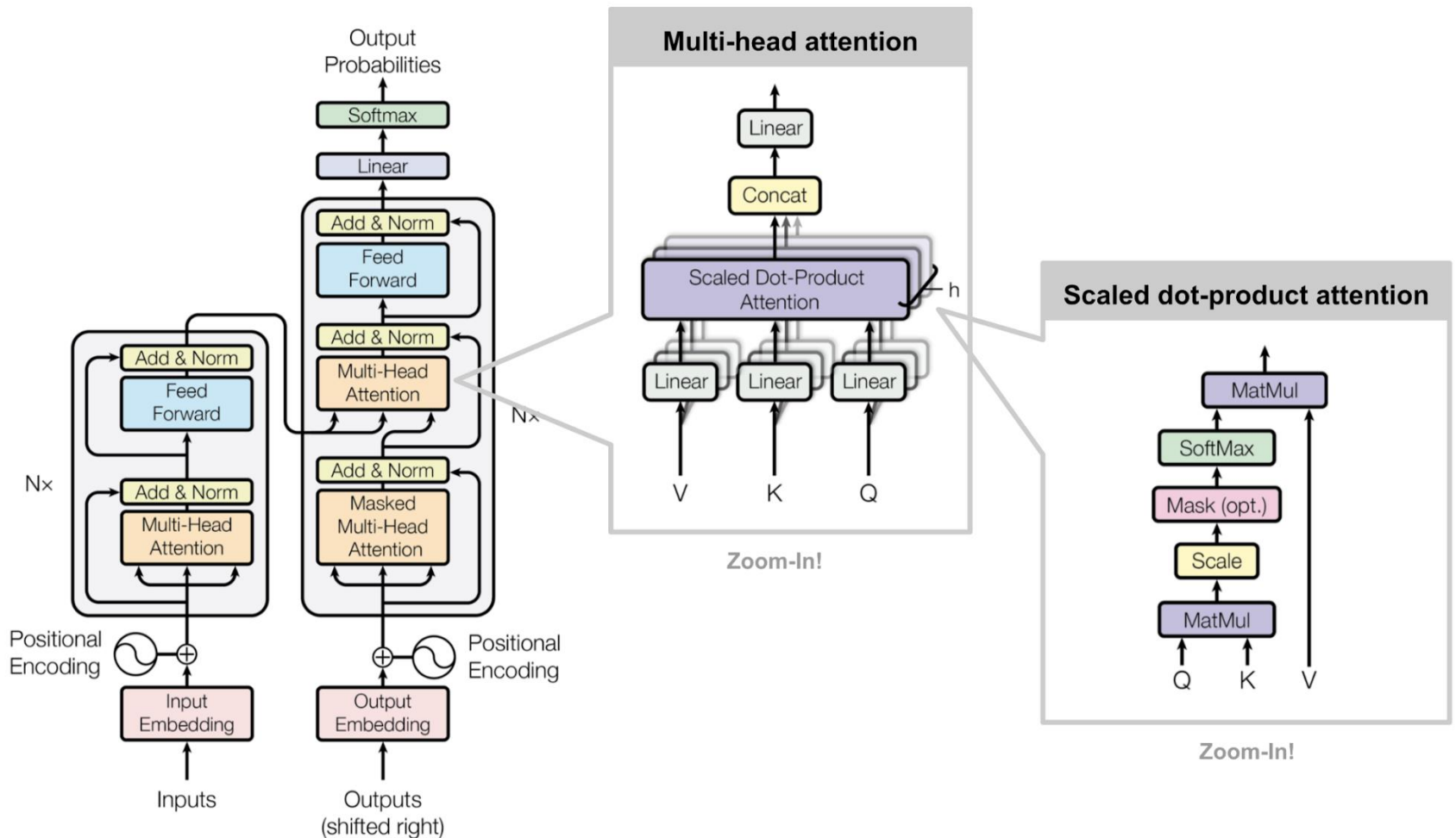
► Multi-head self-attention



Transformer Decoder

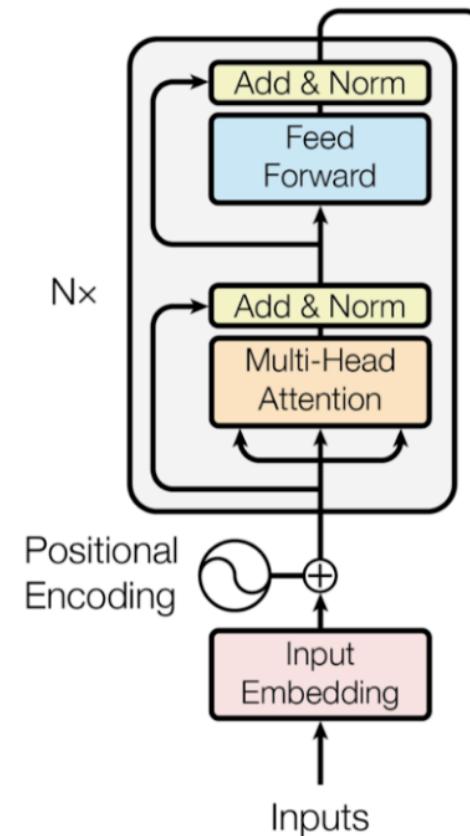


Transformer Architecture Full

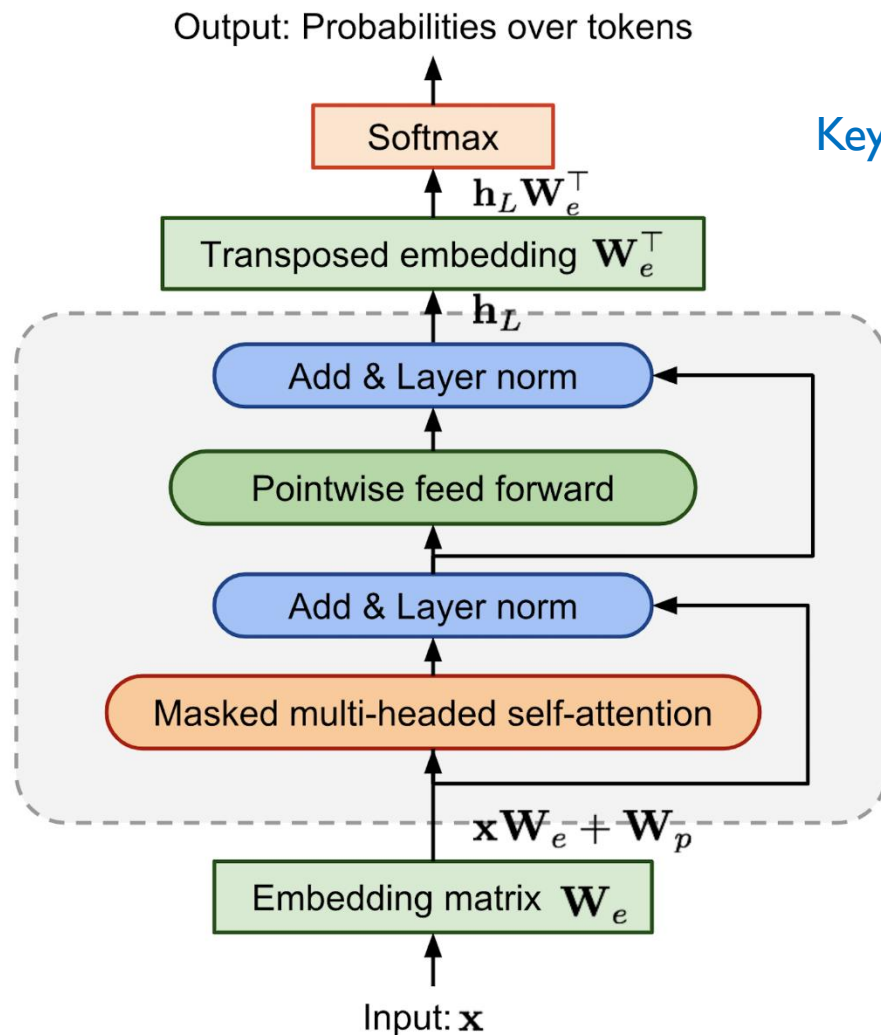


Bidirectional Encoder Representations from Transformers (BERT)

- ▶ Just use Transformer encoder
- ▶ Pre-trained on two tasks to encourage bidirectional prediction and sentence-level understanding
 - ▶ Task 1. Mask language model
 - ▶ Task 2. Next sentence prediction



Generative Pre-training Transformer (GPT)



Key idea: Only use Decoder in Transformer

Transformer Block
Repeat x L=12

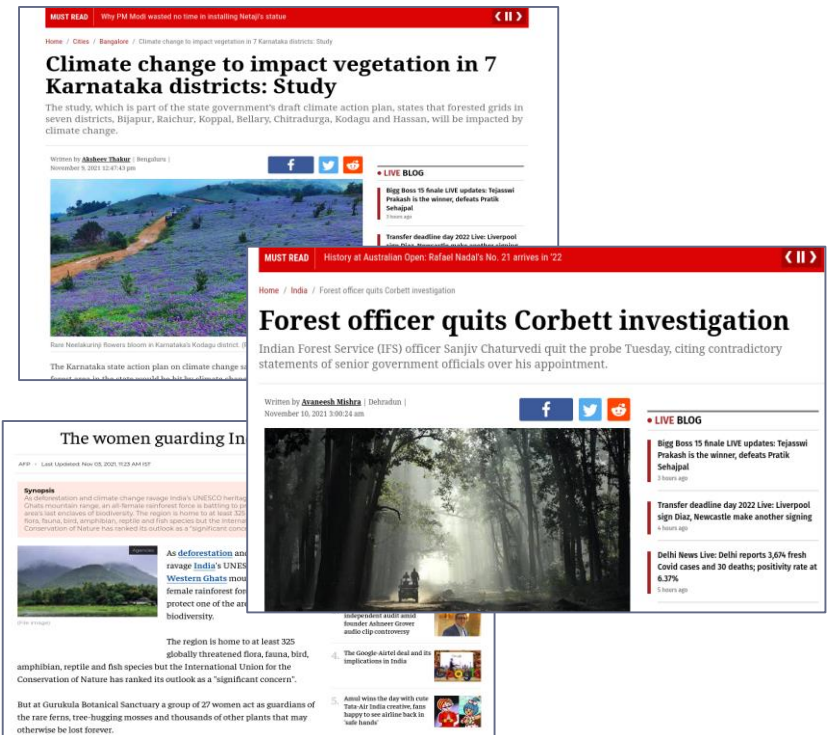
$$\mathbf{h}_\ell = \text{transformer_block}(\mathbf{h}_{\ell-1})$$
$$\ell = 1, \dots, L$$

Outline

- ▶ Transformer revisited
- ▶ NewsPanda
- ▶ NewsSerow
- ▶ Discussion

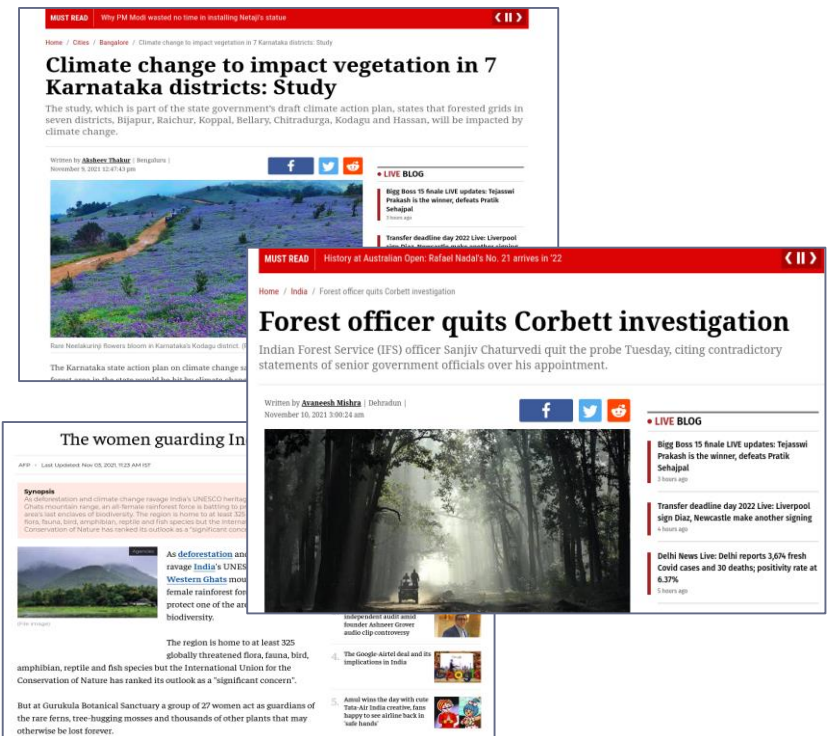
Media Monitoring for Timely Conservation Action

- ▶ WWF country offices spend a lot of time and resources looking through various news articles to identify **trends, events, or threats** related to **conservation and infrastructure**.
 - Identifying 7-10 articles can take 2-3 days to complete



Media Monitoring for Timely Conservation Action

- ▶ Having a tool to **automate this process** will save a lot of time for WWF and allow them to more effectively allocate their resources.
- **How do we identify and analyze media articles for timely conservation and infrastructure actions?**



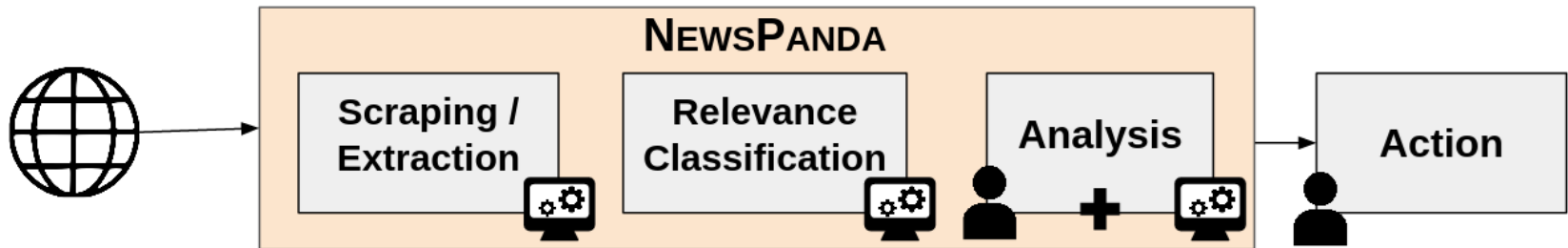
Media Monitoring for Timely Conservation Action

- ▶ Having a tool to **automate this process** will save a lot of time for WWF and allow them to more effectively allocate their resources.
- **How do we identify and analyze media articles for timely conservation and infrastructure actions?**

Infrastructure?

- roads, railways, pipelines, etc.
- Usually high-impact and long-term
- These articles usually cover upcoming developments, which is where WWF can truly perform the necessary interventions

NewsPanda



- ▶ NewsPanda automates multiple steps in the pipeline, enabling humans to perform the more critical tasks (analysis and action)

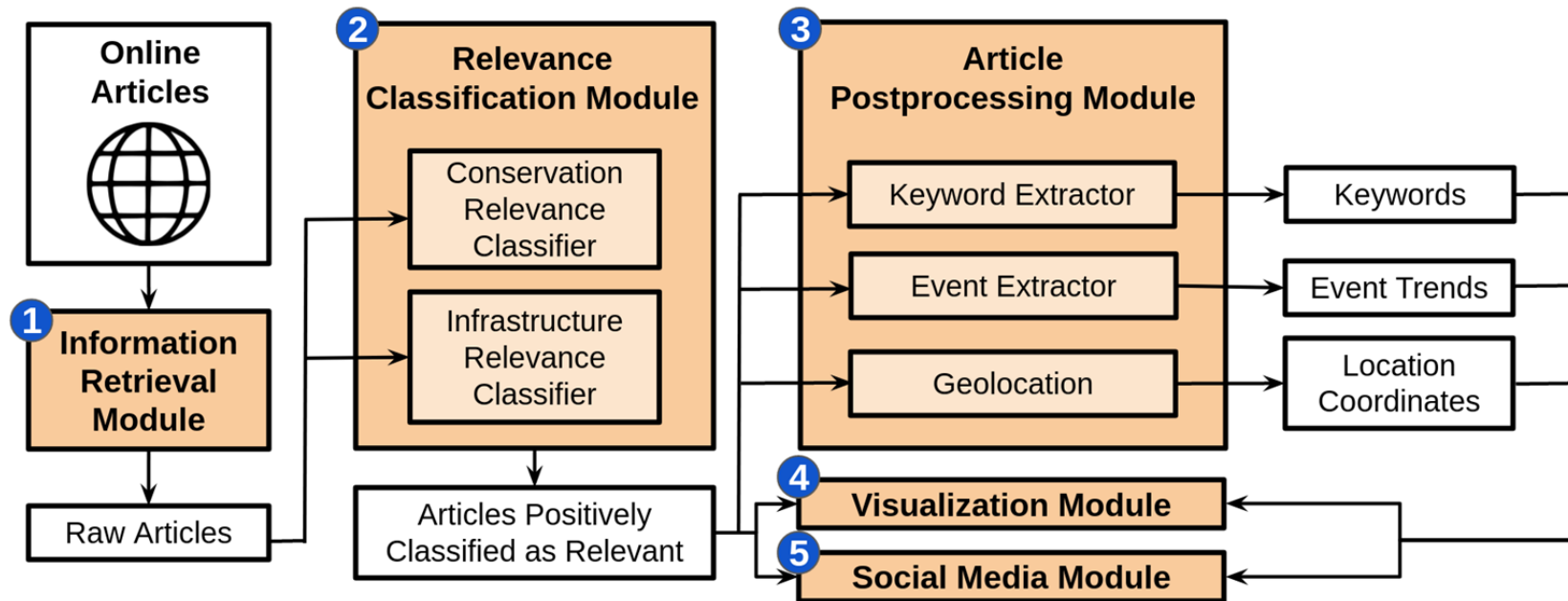
The Team

Practitioners from WWF-UK, WWF-Nepal, WWF-India,
WWF-Norway, WWF-US

Researchers from Carnegie Mellon University



NewsPanda Pipeline



This entire pipeline is ran on a weekly basis.

Dataset

- **Initial dataset** – We start off with two datasets, with labels along two dimensions: conservation relevance and infrastructure relevance
 - **1. WHS-Corp dataset (44,000 articles; 928 with labels)**
 - from Hosseini and Coll Ardanuy (2020)
 - global news articles covering World Heritage Sites around the world
 - only contains labels for conservation relevance

Dataset

- **Initial dataset** – We start off with two datasets, with labels along two dimensions: conservation relevance and infrastructure relevance
 - **1. WHS-Corp dataset (44,000 articles; 928 with labels)**
 - **2. InfraCorp dataset (4,137 articles; 1,000 with labels)**
 - our own dataset which we scrape + annotate
 - focus specifically on India and Nepal
 - scraping done using NewsAPI

Dataset

- **Initial dataset** – We start off with two datasets, with labels along two dimensions: conservation relevance and infrastructure relevance
 - **1. WHS-Corp dataset (44,000 articles; 928 with labels)**
 - **2. InfraCorp dataset (4,137 articles; 1,000 with labels)**
 - each of the 1,000 articles is annotated by two domain experts at WWF
 - Domain experts have limited labeling capacity
 - Q: How do we best select which 1,000 articles out of the 4,137 to label?

Active Learning in Data Collection

- ▶ Confidence-based active learning:
 - ▶ Train an initial model using the available WHS-Corp dataset
 - ▶ Select the 1,000 most “difficult” articles, i.e. the articles which the initial model is “least confident” about

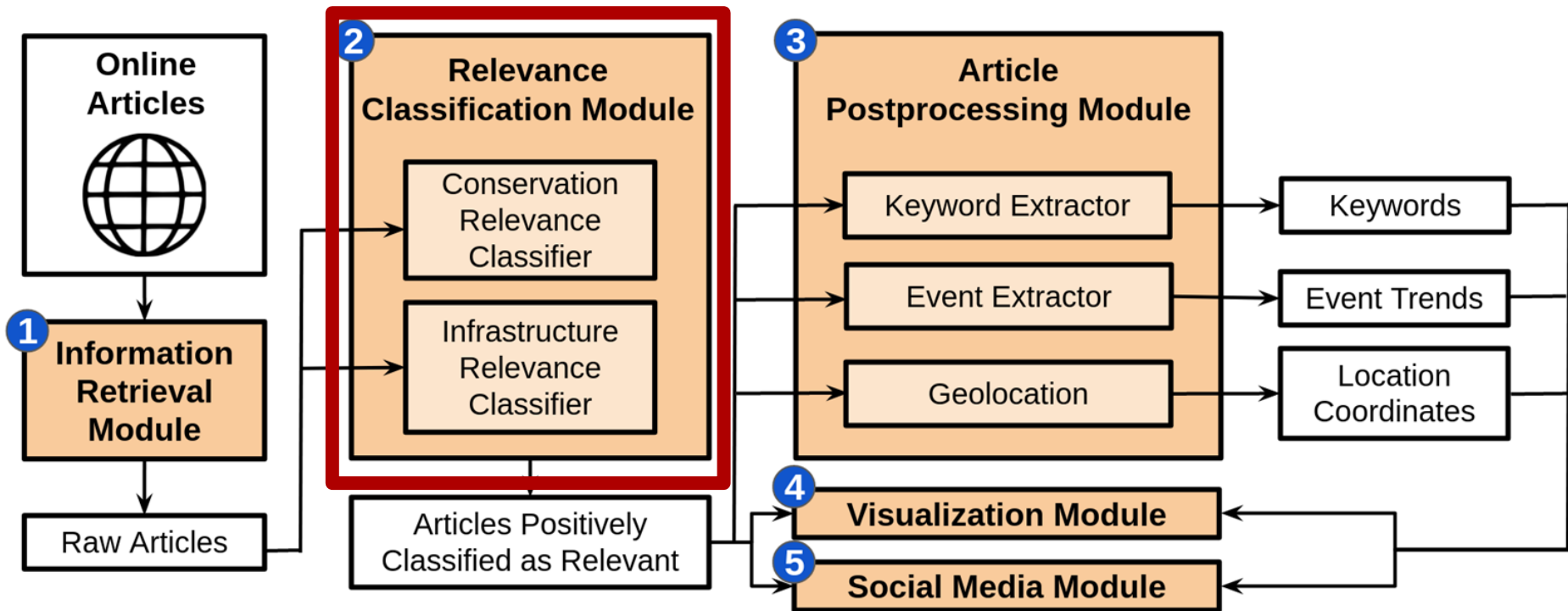
Active Learning in Data Collection

- ▶ Does active learning work? Let's check
 - ▶ Select two sets of 300 articles – set A is actively selected, and set R is randomly selected.

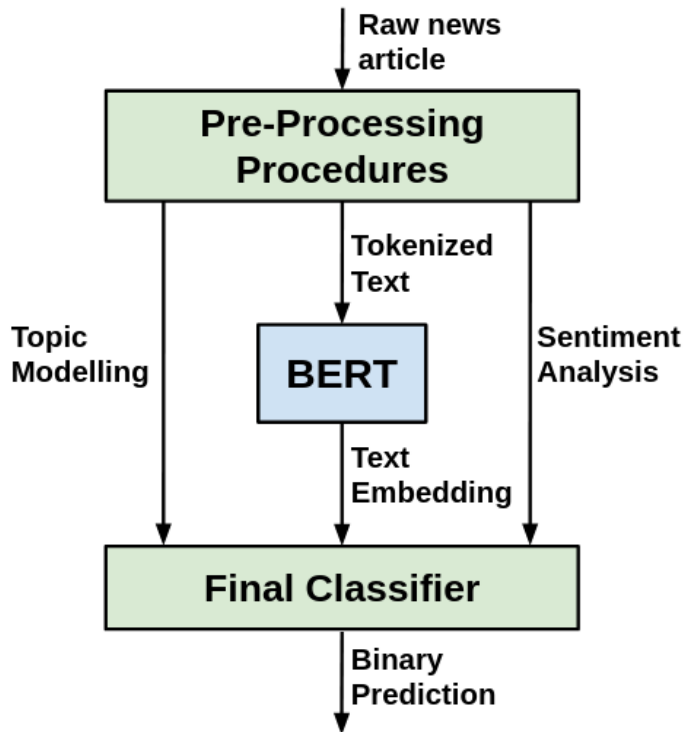
	Dataset	Acc.	P	R	F1
	WHS-CORP	0.911 (0.008)	0.585 (0.035)	0.585 (0.035)	0.586 (0.010)
actively selected	WHS+INF.CORP-A	0.921 (0.004)	0.600 (0.019)	0.774 (0.056)	0.670 (0.019)
randomly selected	WHS+INF.CORP-R	0.916 (0.005)	0.586 (0.035)	0.696 (0.062)	0.637 (0.016)

Using the actively selected set gives a larger performance gain as compared to using a randomly selected set.

NewsPanda pipeline



Relevance Classification Module



We include the following features for each article:

- BERT embedding
- Sentiment analysis score
- Topic modelling vector

Prediction is done along two dimensions: **conservation relevance** and **infrastructure relevance**

Relevance Classification Module

Model	Acc.	P	R	F1
Keyword	0.820 (n/a)	0.317 (n/a)	0.634 (n/a)	0.423 (n/a)
LSTM	0.711 (0.068)	0.495 (0.097)	0.511 (0.129)	0.504 (0.070)
GRU	0.729 (0.054)	0.422 (0.110)	0.505 (0.139)	0.475 (0.067)
BERT	0.860 (0.014)	0.708 (0.032)	0.704 (0.036)	0.706 (0.015)
RoBERTa	0.867 (0.009)	0.705 (0.044)	0.743 (0.041)	0.721 (0.025)
NEWSPANDA	0.877 (0.013)	0.729 (0.032)	0.801 (0.051)	0.744 (0.026)

(a) Scores for *Conservation Relevance*

Model	Acc.	P	R	F1
Keyword	0.947 (n/a)	0.250 (n/a)	0.455 (n/a)	0.323 (n/a)
LSTM	0.908 (0.027)	0.566 (0.160)	0.537 (0.088)	0.554 (0.065)
GRU	0.895 (0.022)	0.544 (0.109)	0.557 (0.123)	0.553 (0.109)
BERT	0.922 (0.018)	0.840 (0.154)	0.745 (0.152)	0.771 (0.096)
RoBERTa	0.916 (0.021)	0.794 (0.091)	0.809 (0.064)	0.799 (0.041)
NEWSPANDA	0.941 (0.018)	0.880 (0.097)	0.821 (0.051)	0.850 (0.043)

(b) Scores for *Infrastructure Relevance*

NewsPanda performs the best across all the models and baselines.

Additional Consideration: Handle Noise Labels

- ▶ Recall the InfraCorp Dataset: each of the 1,000 articles is annotated by two domain experts at WWF
- ▶ How can we handle label noise? Use Noisy label correction methods
 - ▶ Adapt the CORES2 loss (Cheng et al. 2021) noise correction algorithm
 - ▶ Extension of earlier peer loss algorithm – frames the task of learning from noisy labels as a peer prediction problem

$$\ell_{\text{CORES}}(f(x_n), \tilde{y}_n) := \ell(f(x_n), \tilde{y}_n) - \beta \cdot \mathbb{E}_{\mathcal{D}_{\tilde{Y}|\tilde{D}}}[\ell(f(x_n), \tilde{Y})]$$

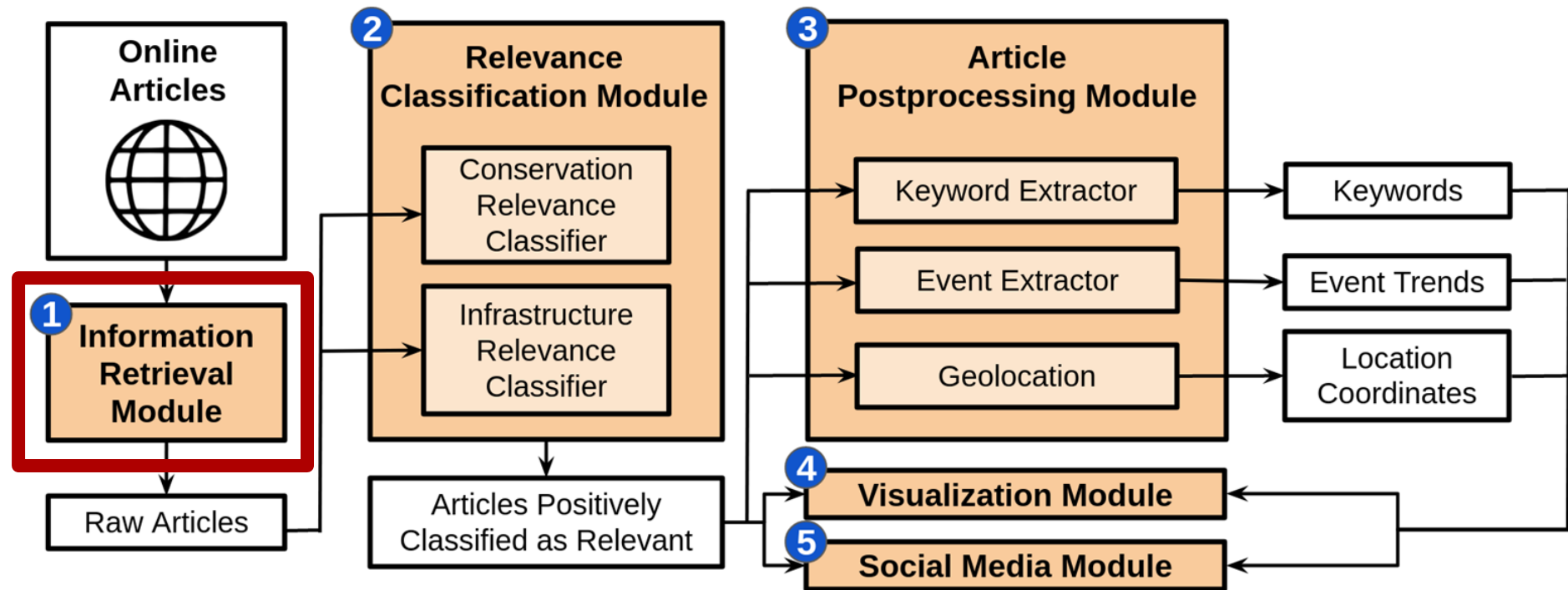
Additional Consideration: Handle Noise Labels

- ▶ Does the noisy label correction algorithm?
- ▶ Investigate the effects of using peer loss and CORES2 loss

Noisy Label Correction	Acc.	P	R	F1
None	0.907 (0.004)	0.566 (0.015)	0.441 (0.055)	0.497 (0.026)
Peer Loss	0.911 (0.006)	0.591 (0.031)	0.465 (0.027)	0.509 (0.017)
CORES ²	0.908 (0.009)	0.584 (0.057)	0.551 (0.050)	0.553 (0.014)

Using CORES² loss yields the best performance compared to using Peer Loss and using no noisy label correction.

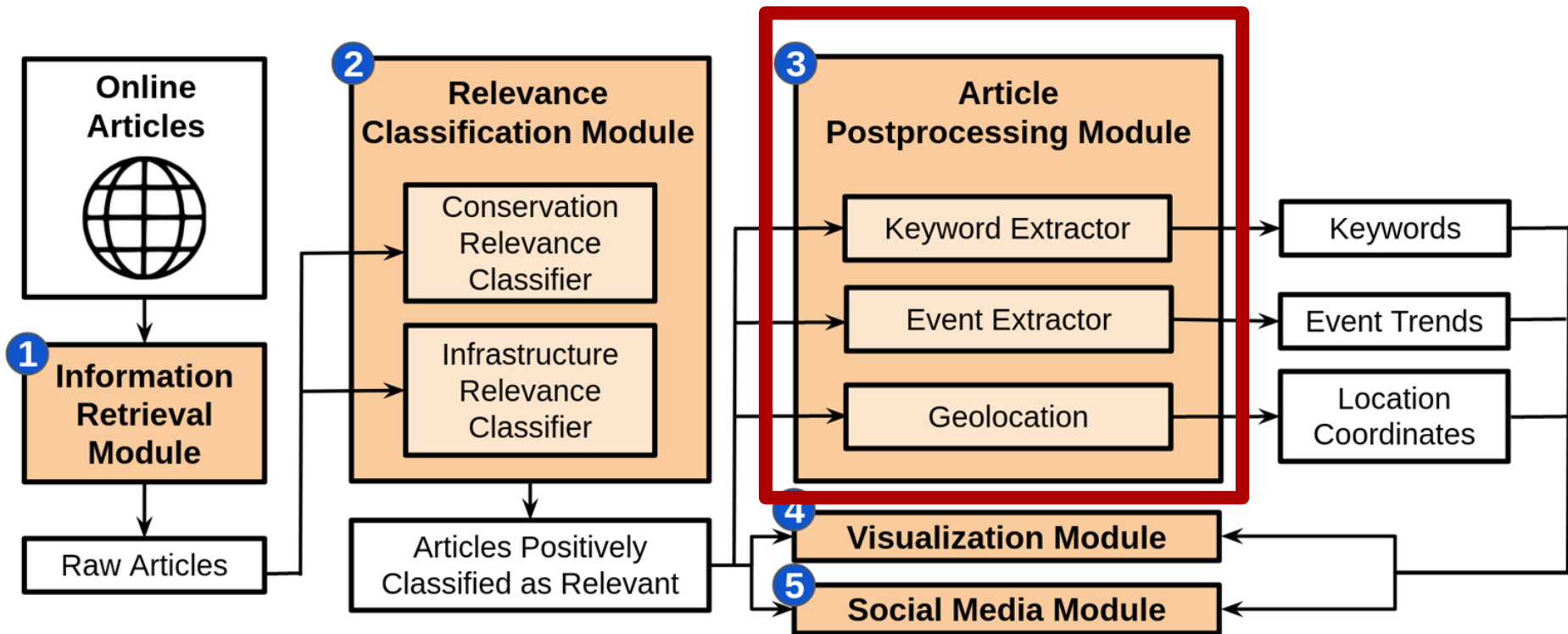
NewsPanda pipeline



Information Retrieval Module

- ▶ Use the NewsAPI scraper with search terms taken from a list of curated conservation sites by WWF
- ▶ Focus on India and Nepal
- ▶ This is ran on a weekly basis

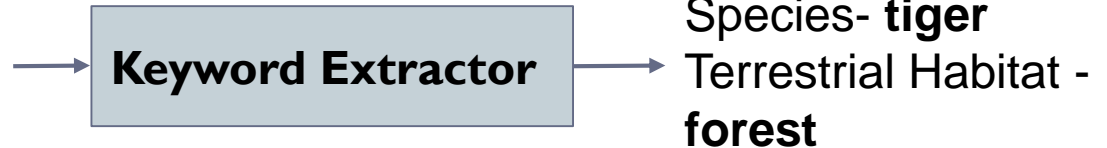
NewsPanda pipeline



Keyword extraction

(example paragraph)

A 45-year-old man from Chamrajanagar in Karnataka was arrested by the Forest Department for attempting to hunt wild animals and also setting fire in **forest** areas in the Talavadi Forest Range in the Sathyamangalam **Tiger** Reserve here.



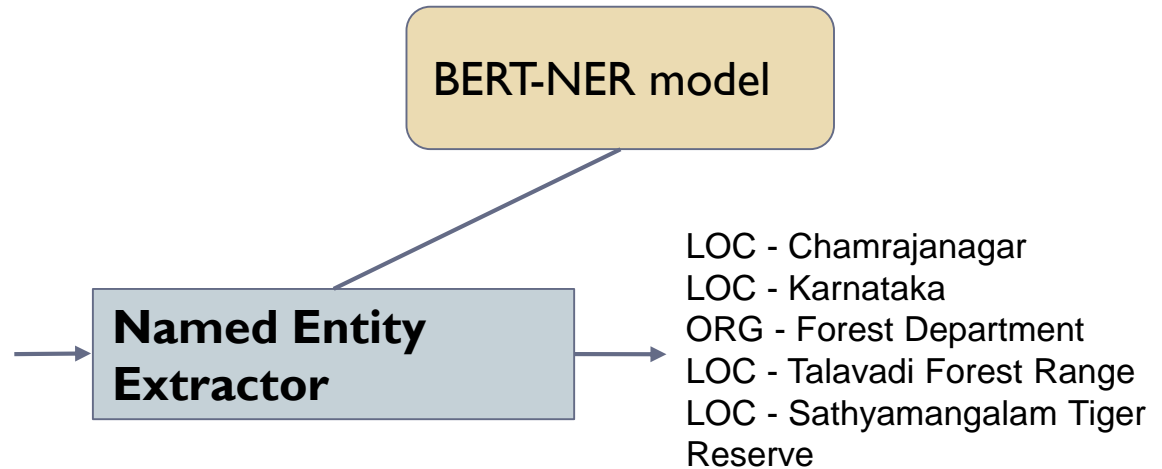
	A	B
1	Atmosphere	carbon
2	Atmosphere	CH4
3	Atmosphere	CO2
4	Atmosphere	methane
5	Atmosphere	nitrogen
6	Conservation Keyword	adaptation
7	Conservation Keyword	adaptive management
8	Conservation Keyword	alternative livelihoods
9	Conservation Keyword	animal identification
10	Conservation Keyword	animal traceability
11	Conservation Keyword	animal welfare
12	Conservation Keyword	benefit sharing

We use a list of around 1000 keywords (level 1 and level 2), then check for matches in the text.

Named entity recognition

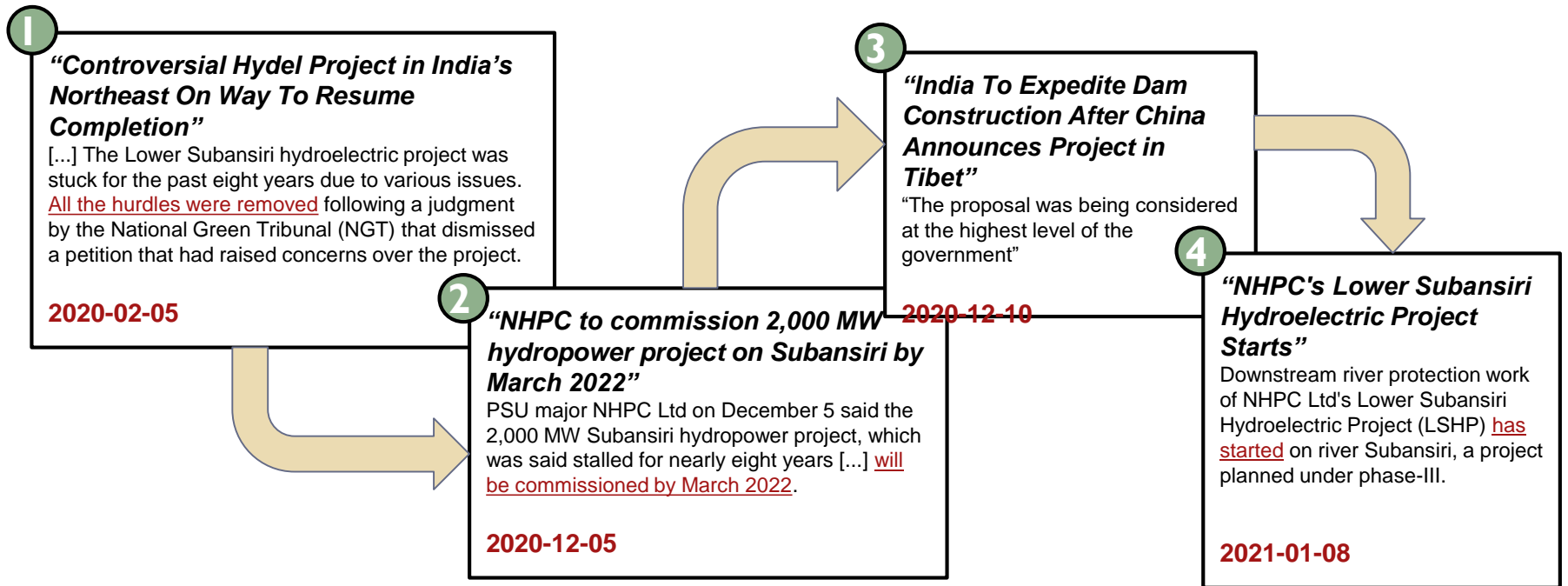
(example paragraph)

A 45-year-old man from **Chamrajanagar** in **Karnataka** was arrested by the **Forest Department** for attempting to hunt wild animals and also setting fire in forest areas in the **Talavadi Forest Range** in the **Sathyamangalam Tiger Reserve** here.



Event timeline

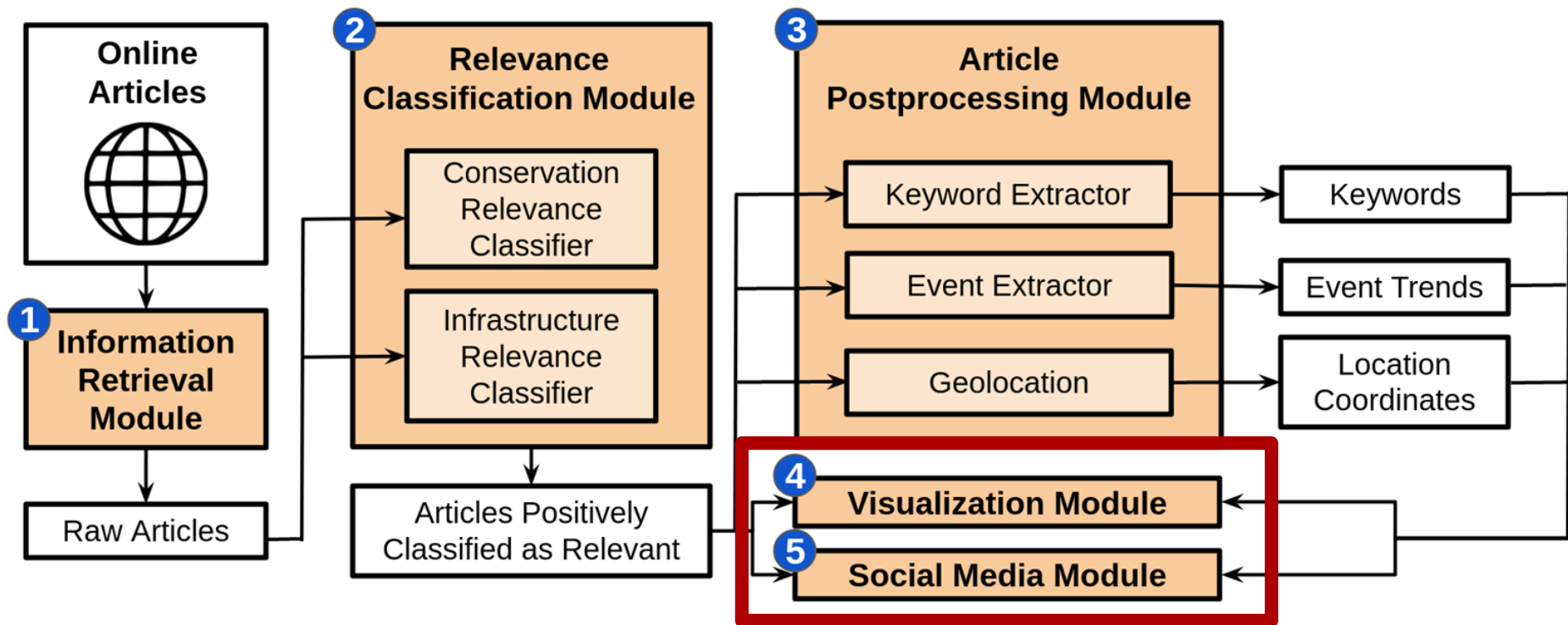
We search for entity “**Subansiri**”, then filter with the keyword “**hydroelectric**”.



Geolocation

- ▶ Important to integrate into WWF's GIS systems
- ▶ Use a directory of conservation sites from WWF to map articles to their coordinates
- ▶ If there is no match in directory, we use the geopy package

NewsPanda pipeline



Deployment

- ▶ NewsPanda has been deployed by WWF teams in India, Nepal, and the UK since February 2022
- ▶ Three stages of deployment:
 - ▶ Pilot study (February 2022)
 - ▶ Goal: Test out the pipeline and identify some operational and technical issues in the initial version of NewsPanda
 - ▶ Initial deployment (March 2022 to July 2022)
 - ▶ Goal: Evaluate the performance of NewsPanda
 - ▶ Sustainable deployment (August 2022 onwards)
 - ▶ Goal: Make pipeline more automatic and cloud-based

Deployment Results

► Quantitative results:

- Each week, the WWF teams from India, Nepal, and the UK evaluated the articles classified by NewsPanda

	Conservation			Infrastructure		
	P	R	F1	P	R	F1
WWF India	0.849	0.605	0.706	0.462	0.250	0.324
WWF Nepal	0.895	0.917	0.906	0.923	0.308	0.462
WWF UK	0.879	0.823	0.850	1.000	0.455	0.625

- **High precision values** = trustworthy and reliable system
- **Low recall for infrastructure** = misses out on potential articles; needs to improve more on positively identifying relevant articles

Deployment Results

► Qualitative results:

- Two months into deployment, the CMU team carried out semi-structured interviews with their WWF colleagues who have been using NewsPanda outputs in their work

“You’re giving us a bunch of articles... over 50 articles a week. We had two interns who spend 2-3 days a week on this and would only give us seven to ten articles. So there is a huge bump in efficiency right there in itself.”

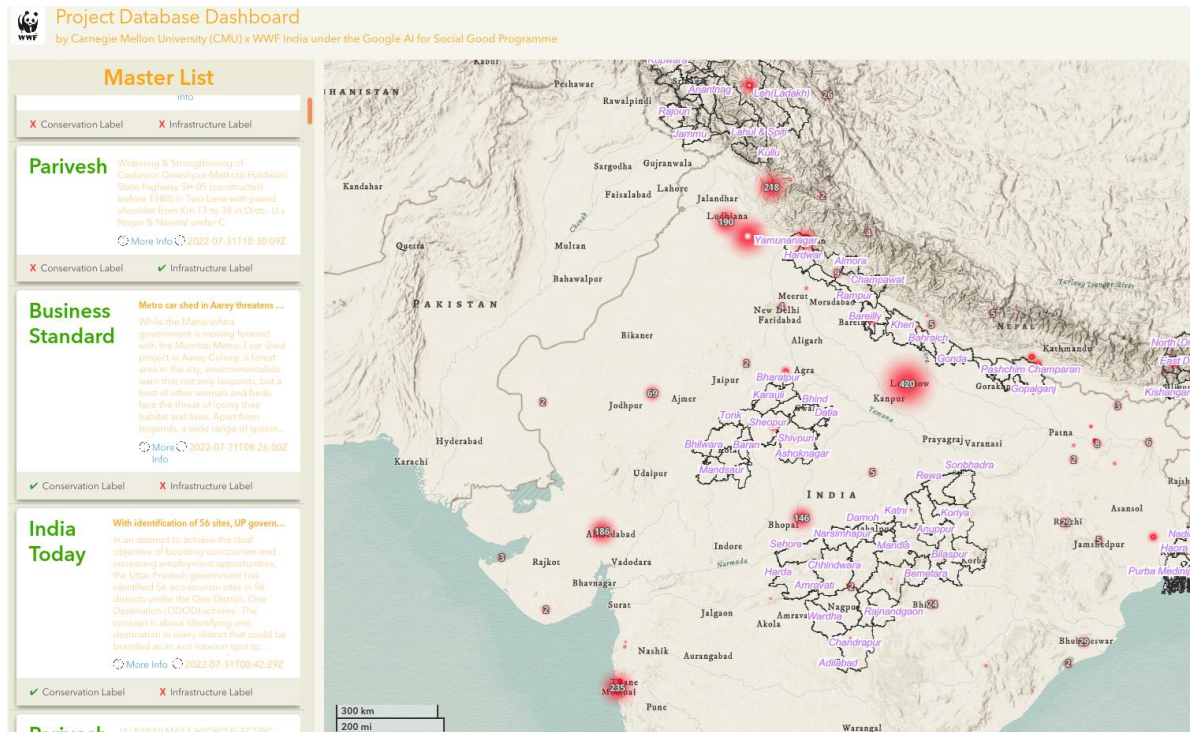
“It took us maybe a month to do analyses of three or four infrastructure projects. With **NEWSPANDA**, we can send (stakeholders) 20 or 30 reports in a month.”

“The data that you’re sharing give a global perspective. It is very useful to understand the upcoming projects or mitigation measures that are being adopted on a global scale. So it helps us be informed.”

“It’s also a transition in their (WWF staff) job function. They will not just be doing data hunting. They are qualifying themselves to be data analysts.”

Visualization Module

- ▶ The NewsPanda results are integrated into WWF's GIS systems, which is especially useful for the field teams.

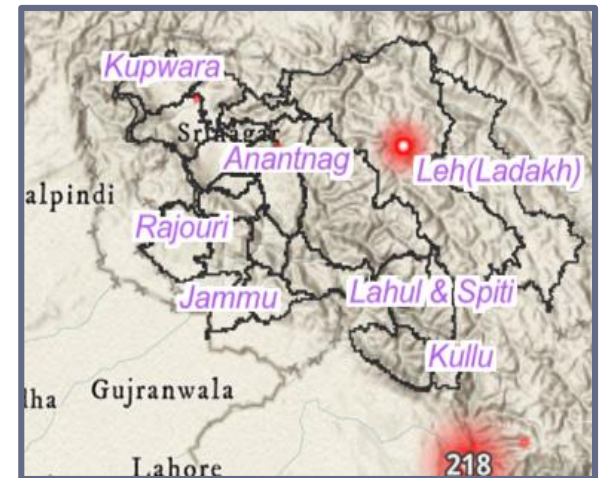


Visualization Module – Success Story

August 2022: NewsPanda highlighted **Ikhala Block Boundary Kishtwar to Lopara Road** in the GIS system

Upon further investigation, it is found that the project would **divert 5.9 hectares of forest land**

More importantly, WWF found that the project was still at its **pre-proposal stage**. This means WWF would be able to take early action and possibly participate in relevant conversations.



Social Media Module

For the general public to benefit from NewsPanda, we also developed a Twitter bot which tweets links and hashtags (keywords) to the relevant weekly articles.

@WildlifeNewsIND

Go follow and share! :)



Lessons Learned

- ▶ Quick prototyping. Simple models, rapid iterations, discover new problems.
- ▶ The “not-so-AI” components realize the promise of AI for nonprofit project.
- ▶ Nonprofit ownership. No fancy dashboards that nobody use.
- ▶ Get feedback, in formal form.



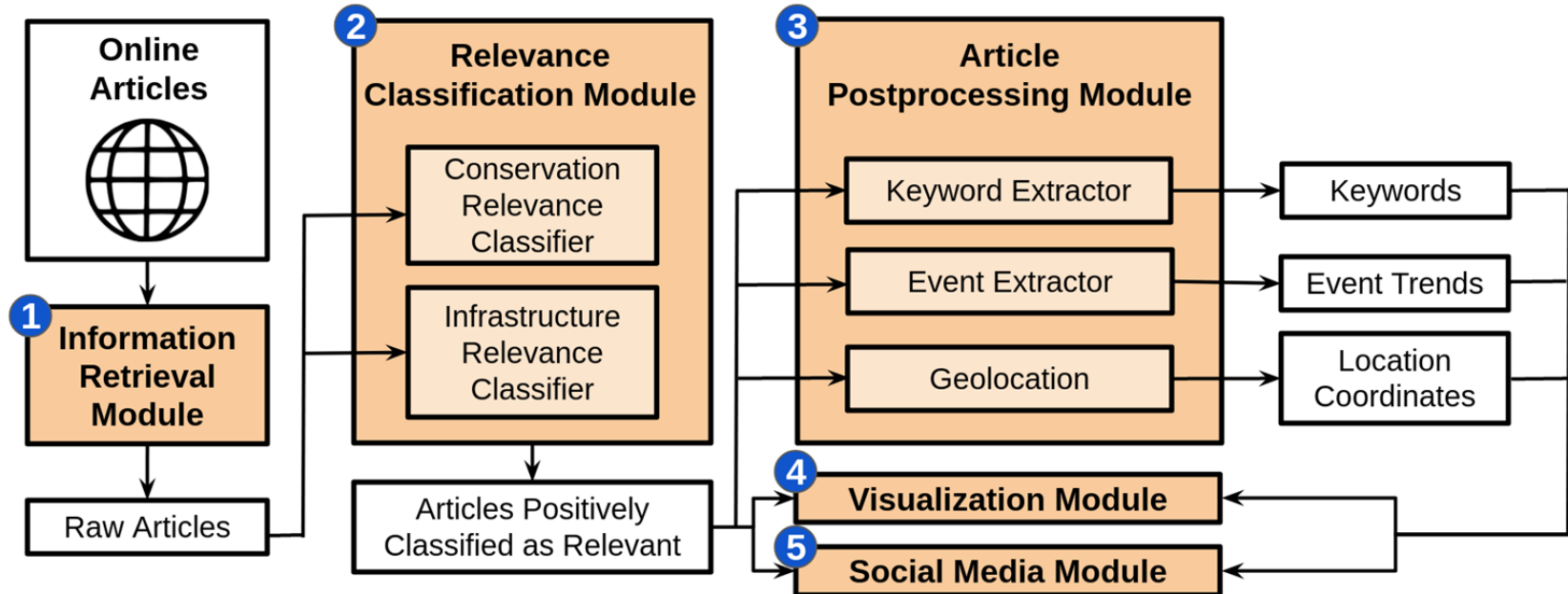
Poll I

- ▶ Which of the following ones are used in the relevance classification module of NewsPanda?
 - ▶ A: Topic modeling
 - ▶ B: Sentiment analysis
 - ▶ C: BERT embedding
 - ▶ D: ChatGPT

Outline

- ▶ Transformer revisited
- ▶ NewsPanda
- ▶ NewsSerow
- ▶ Discussion

Recap: NewsPanda Pipeline



Classification module trained on two datasets

- 1. WHS-Corp dataset (44,000 articles; 928 with labels)
- 2. InfraCorp dataset (4,137 articles; 1,000 with labels)

Limitation: Only contains English Articles

Extensions to NewsPanda

- ▶ Extend NewsPanda to handle multiple languages, including low-resourced languages, e.g., Nepali
- ▶ Challenge: Hard to extend NewsPanda in a scalable fashion if we need to collect >1000 labeled articles for each new language/country

Translation Test

- ▶ Step 1: Translate the news article in the test set from Nepali to English using existing translation tools (e.g., Google translate)
- ▶ Step 2: Use the relevance classification module in NewsPanda to classify the translated test article
- ▶ Pro: Does not need any labeled data in Nepali
- ▶ Con: If we do have some labeled data in Nepali, this workflow does not make good use of it

Translation Train

- ▶ Step 1: Translate the news articles in the training set from Nepali to English using existing translation tools (e.g., Google translate)
- ▶ Step 2: Fine-tune the relevance classification module in NewsPanda using the translated articles and corresponding labels
- ▶ Step 3: Translate the test article into English
- ▶ Step 4: Use the relevance classification module in NewsPanda to classify the translated test article

Extensions to NewsPanda

- ▶ Can we use ChatGPT to tackle the challenge with low resource languages directly?

Attempt 1

Model: Default (GPT-3.5)



Is this text relevant to environmental conservation?

वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा



I'm sorry, but the text you provided is written in the Nepali language and does not appear to relate to environmental conservation. Can you please provide a text in English or any other language that I can help you with?



Attempt 2: Restrict the answer space

Model: Default (GPT-3.5)

LE

Is the following text relevant to environmental conservation?

वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा

A) Yes. B) No.



B) No.



The text appears to be in Nepali language and does not relate to environmental conservation.



Attempt 3: Get a sense of uncertainty

Model: Default (GPT-3.5)

LE

The following text contains a Nepali news headline. On a continuous scale from 0.00 to 1.00, how related is the headline to environmental conservation?

वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा

Conservation score:



Without any further context, it is difficult to assess the exact relevance of the Nepali news headline to environmental conservation. However, based solely on the headline, it does not appear to have any direct connection to environmental conservation. Therefore, the conservation score would be 0.00.



Attempt 4: Use GPT-4

Model: GPT-4

LE

The following text contains a Nepali news headline. On a continuous scale from 0.00 to 1.00, how related is the headline to environmental conservation?

वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा
Conservation score:



वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा



This headline, translated to English, reads: "Second climbing permit granted in the spring season for Manaslu."

Conservation score: 0.35

This headline is somewhat related to environmental conservation as it discusses climbing permits in the Manaslu region, which could potentially have an impact on the environment. However, it does not directly address conservation efforts or concerns.



Attempt 5: Show ChatGPT some examples (In-Context Learning)

Model: Default (GPT-3.5)

LE

Each of the following examples contains a Nepali news headline and label classifying whether or not it is related to environmental conservation. Observe the examples and fill in the last label:

अपडेट : लिदी पहिरोमा पुरिएका २ जनाको शव भेटियो
Conservation label: No

वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा
Conservation label: No

पर्यटनले मुहार फेरेको भकुण्डे गाउँ
Conservation label: Yes

वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा
Conservation label:



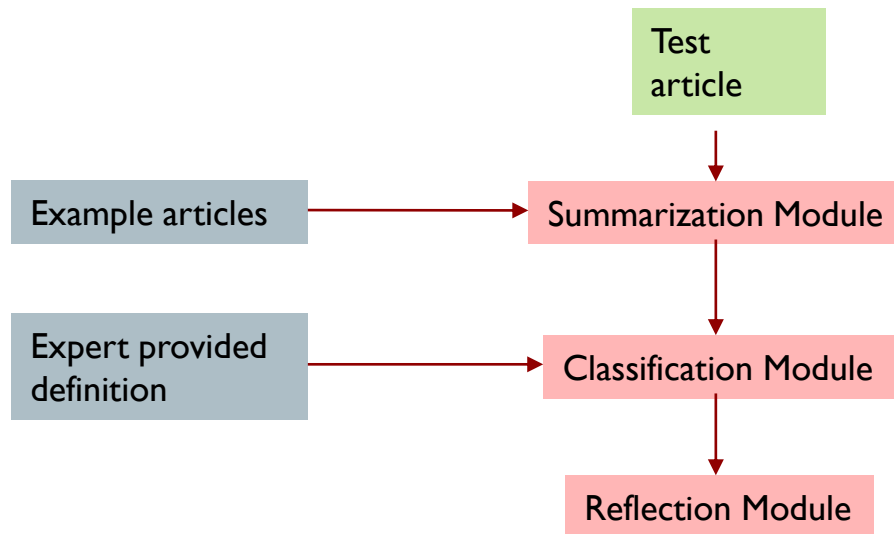
No



Discussion

- ▶ Discuss with your neighbor what prompt you would like to try to interact with ChatGPT to get the conservation relevance label for a news article with the following title:
- ▶ वसन्त ऋतुमा दोस्रो आरोहण अनुमति मनास्लुमा

► Summary + In-Context Learning + Reflection



Evaluation

- ▶ With only a few examples, NewsSerow achieves comparable performance to fine-tuned models (which uses much more training data) for Nepali articles

Model	Nepal/Nepali		
	Precision	Recall	F1-Score
Zero-shot Models			
GPT-3.5-Turbo	0.79 (0.02)	0.32 (0.01)	0.46 (0.01)
Few-shot Models			
mBERT	0.26 (0.03)	0.70 (0.12)	0.37 (0.01)
XLNet	0.35 (0.04)	0.60 (0.23)	0.43 (0.06)
GPT-3.5-Turbo	0.68 (0.05)	0.58 (0.07)	0.62 (0.02)
NewsSerow	<u>0.88</u> (0.03)	<u>0.58</u> (0.01)	0.70 (0.01)
Fine-tuned Models			
mBERT	0.77 (0.06)	0.57 (0.05)	0.65 (0.02)
XLNet	0.70 (0.05)	0.71 (0.04)	0.70 (0.01)
Translation Test	0.74 (0.16)	0.64 (0.15)	0.66 (0.06)

Deployment

- ▶ NewsSerow has been deployed at WWF offices in Nepal and Colombia since Apr 2023 (latest version in use since Aug 2023)
- ▶ Results from 8 weeks of deployment:

Week	Nepal/Nepali				Colombia/Spanish			
	# Ex.	P	R	F1	# Ex.	P	R	F1
1	10	0.5	1.00	0.67	28	0.57	1.00	0.73
2	9	0.00	0.00	0.00	22	1.00	0.70	0.82
3	21	1.00	0.33	0.50	28	0.5	0.8	0.62
4	9	1.00	0.75	0.86	26	0.63	0.71	0.67
5	9	0.50	0.50	0.50	35	0.67	0.80	0.73
6	5	0.50	1.00	0.67	23	0.38	0.50	0.43
7	8	1.00	1.00	1.00				
8	13	1.00	0.38	0.55				
Aggr.	84	0.77	0.55	0.64	162	0.61	0.73	0.67

Discussion

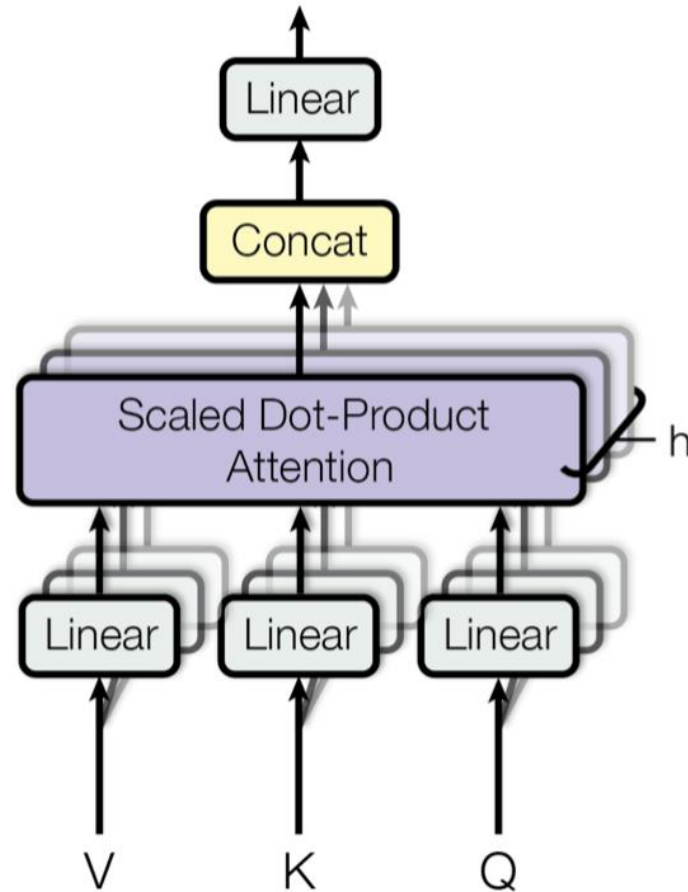
- ▶ Can we build NewsPanda or NewsSerow-like systems to help achieve other UN SDGs?



Backup Slides

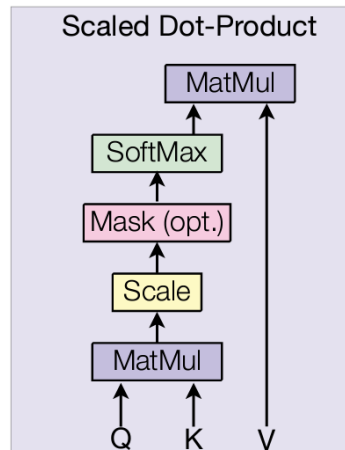
Transformer

► Multi-head self-attention



Scaled Dot-Product Attention in Transformer

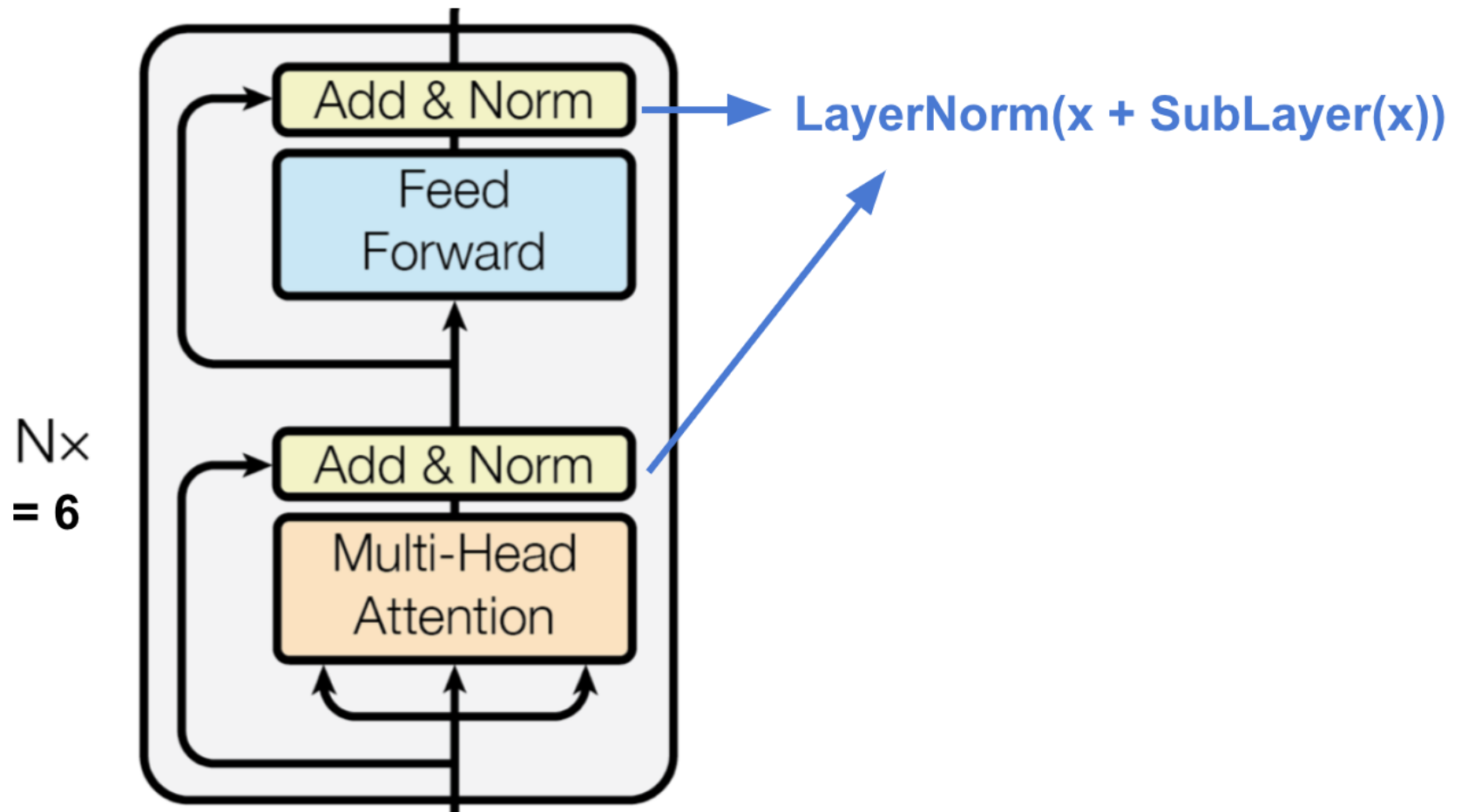
$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V$$



The diagram illustrates the matrix operations for Scaled Dot-Product Attention. It shows the input matrices Q (purple, 2x3) and K^T (orange, 3x3) being multiplied together (×). The result is divided by the square root of the dimension d_k (√d_k). The result is then passed through a softmax function. The final result is multiplied by the input matrix V (blue, 2x3) to produce the output matrix Z (pink, 2x3).

$$\text{softmax}\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) V = Z$$

Transformer Encoder



Transformer Decoder

