

Live Reality: Up-to-Date Information Quality



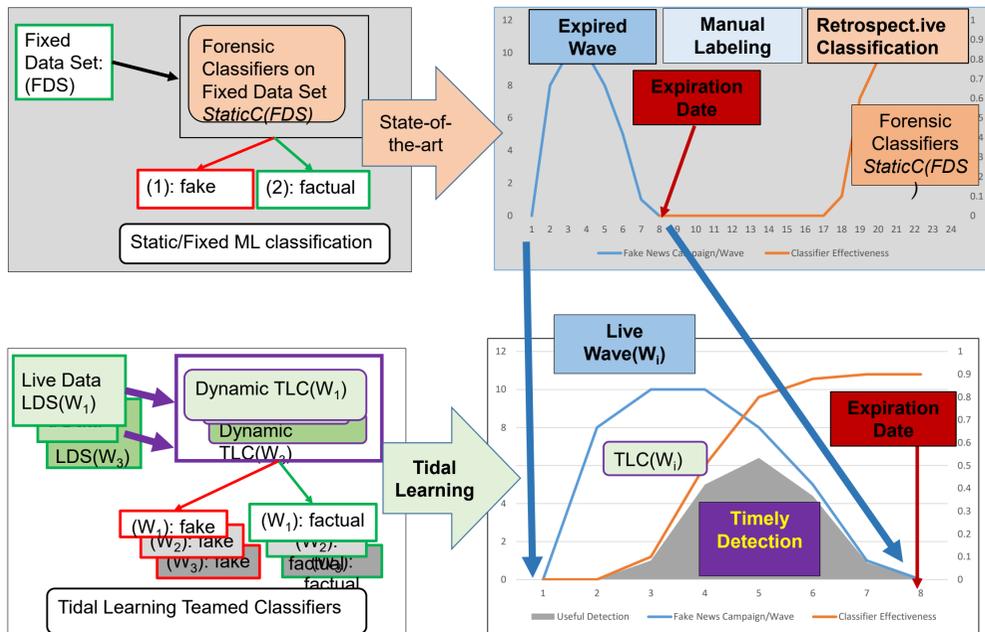
PI: Calton Pu, School of Computer Science, Georgia Institute of Technology

Project URL: <https://grait-dm.gatech.edu/>

Vision: Detection and Filtering of Fake News during Their Dissemination

Challenges:

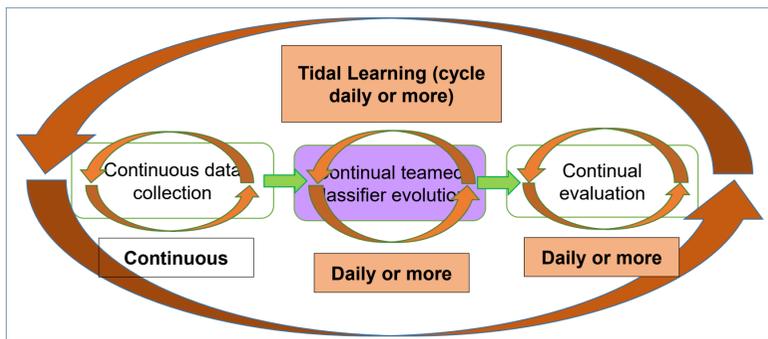
- Retrospective (forensic) studies of fake news are *old news*. Each new wave of fake news is different by construction of “breaking news”.
- Authoritative sources of ground truth are domain-specific (e.g., CDC for the COVID-19 pandemic).
- New software tools for automating continuous data collection, teamed classifier updates, and their evaluation.



State-of-the-art: forensic studies of manually labeled fake news after a wave has expired



What we need: detect (and filter) new fake news waves while they are still raging

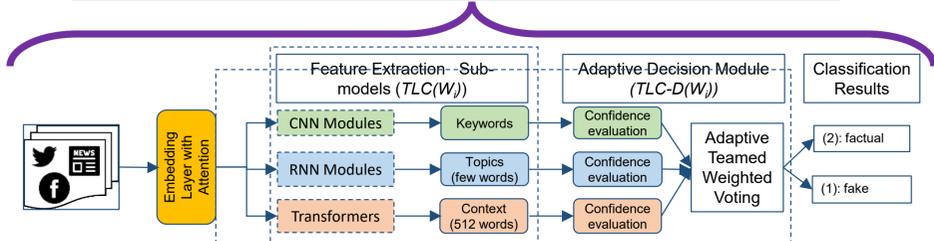


Tidal Learning (training cycle):

- Continuous collection of training and test data
- Continual evolution of live teamed classifier (from new training data)
- Continual evaluation of live teamed classifier (by new test data)

Scientific Impact:

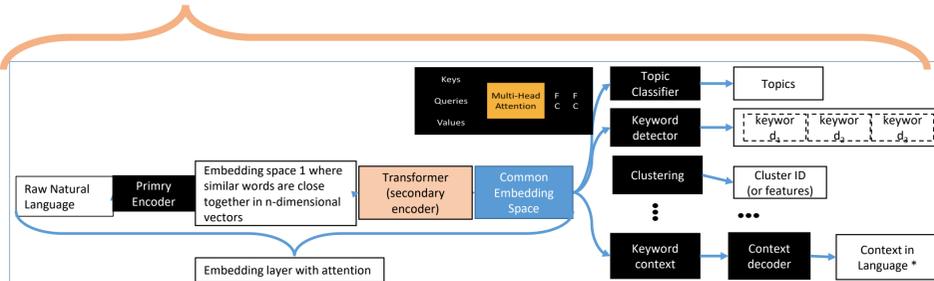
- First attempt to identify and filter fake news in real-time, in daily cycles as soon as they are created
- Study of live fake news waves through evolution of authoritative and test data
- Moving beyond the (past) gold standard of retrospective/forensic studies based on manual labeling



Automating the process of creating/updating new sub-models and decision module during the daily training cycle

Broader Impacts and Broader Participation:

- Software tools for filtering new fake news at the same efficiency and timeliness as email spam.
- Relief and reduction of harm caused by infodemic.
- Sharing of new fake news data and authoritative (training) data in daily cycles as they are created.
- Sharing of live classifiers and their evaluation results!
- Sharing of software tools for continuous data collection, classifier training, and evaluation experiments.



Automating the process of embedding (keywords and context) during the daily training cycle

Live Fake News Wave (W_i), Days 1 through 8					Expired Fake News: Days 9 through 24	
Period	Day 1 (p_1)	Day 2 (p_2)	...	Day $n=8$ (p_n)	Day 9: End of wave, (Expiration Date)	(State-of-the-art) Retrospective classification by <i>StaticC(FDS)</i> , corresponding to one square in the table
Team class.	$TLC(p_1)$	$TLC(p_2)$...	$TLC(p_n)$		
Evaluation	$TLC(p_1)$ on $LDS-T(p_1)$	$TLC(p_2)$ on $LDS-T(p_1)$...	$TLC(p_n)$ on $LDS-T(p_1)$	Manual Labeling	
Evaluation	$TLC(p_1)$ on $LDS-T(p_2)$	$TLC(p_2)$ on $LDS-T(p_2)$...	$TLC(p_n)$ on $LDS-T(p_2)$		
...		
Evaluation	$TLC(p_1)$ on $LDS-T(p_n)$	$TLC(p_2)$ on $LDS-T(p_n)$...	$TLC(p_n)$ on $LDS-T(p_n)$	Tidal learning: Timely (daily) classification	Tidal learning: Retrospective classification

Experiment Design (Part 1: during the live wave, days 1 through 8, left side of table)

- Create new sub-models to add to teamed classifier
- Update adaptive decision module to reflect true novelty
- Evaluate the updated teamed classifier

Experiment Design (Part 2: after expiration, days 9 and later)

- Retrospective studies of teamed classifier at various moments in time
- Evaluation of Tidal Learning approach in timeliness and variations in accuracy

Tidal Learning live classifiers: created and evaluated in daily cycles

Forensic classifiers