

31st ACM International Conference on Information and **Knowledge Management**

October 17-21, 2022

CIKM2022

Hybrid Conference, Hosted in Atlanta, Georgia, USA

DISCO: Comprehensive and Explainable Disinformation Detection



Dongqi Fu (UIUC)



Yikun Ban (UIUC)



(UIUC)



Hanghang Tong Ross Maciejewski (ASU)



Jingrui He (UIUC)

Presenter: Dongqi Fu

Email: dongqif2@illinois.edu









Motivation

- Disinformation (e.g., fake news) is fabricated to mislead the general public. When we facing a piece of suspicious information,
 - Is a piece of suspicious information fake or real?
 - How to deal with polysemy in this piece of suspicious information?
 - Which words, phrases, or sentences are dominant in determining this information fake or not?
 Real or Fake?



Disinformation Detection and Explanation: DICSO

- Input: A suspicious news article
- Output:
 - (1)The fake news probability, (2) Misleading degree of each word
- Techniques and Architecture
 - (1)Word graph construction, (2) Geometric feature extraction
 - (3)Neural detection, (4) Misleading degree analysis
- Performance

~97.5% precision and recall on a public dataset [1] Building Send DISCO: Comprehensive and Explainable Disinformation Detection Word Enter your articles Graph Californians have a choice ahead of them. In November, citizens of the Golden State will decide whether or not to - Building the Word Graph for the Input Article legalize marijuana for recreational purposes. One of the most compelling arguments proponents of marijuana --- Extracting Geometric Features legalization have used in previous efforts to legalize marijuana in other states, is that it will bring in extra money --- Making Predictions for the state s budget. On that front, there is plenty of evidence to support that argument. Oregon saw a --- Analyzing Each Word Misleading Degree Geometric jaw-dropping amount of extra revenue after they legalized marijuana. Officials in Oregon expected to bring in \$2 --- DISCO Finished Input Article from User's Query: million dollars in tax revenue after their first year of legal weed. They nearly doubled their expectation bringing Feature in an astounding \$3.48 million dollars in new tax revenue. So how much is California expected to bring in should Working Flow "Taking a hot bath prevents the COVID-19 they choose to legalize recreational marijuana? Reuters reports that California could take in an additional \$1 Extraction billion dollars a year: With a population of nearly 40 million people, and a thriving medical marijuana trade disease." legalized 20 years ago, California already has the United States largest legal marijuana market. Legalization of Prediction recreational pot would generate an estimated \$1 billion in additional taxes per year. Other states that have legalized recreational marijuana have received similar tax windfalls. Colorado, the first state in the U.S. to Detection Result: Fake legalize marijuana, took in \$135 million dollars in taxes and fees in 2015. Colorado is pumping most of their new **Output Results:** Neural Fake Probability: 0.99999998466798 Credibility tax dollars into schools and other education initiatives. The state of Washington raised \$67.5 million dollars in Real Probability: 0.00000001533202 tax revenue from recreational marijuana in 2015. They are expected to raise more than \$1 billion dollars over (1) Detection Decision and Feature Vector Detection [Word]: [Misleading Degree] four years. This money goes into Washington's general fund and to help fund health services. Legalizing marijuana recreational: 0.00000000949085 (2) Misleading Words Ranking would also save California a lot of money due to no longer having to pay the enormous cost of enforcing current 0.00000000810895 marijuana laws. In 2012, over 21,000 people were arrested for crimes related to marijuana criminalization. The legalize: 0.0000000645399 number of overall arrests especially felony arrests has plummeted over the years after changes were made to the Word's previous: 0.00000000564683 California s marijuana laws. However, the cost both budgetary and human of marijuana criminalization is still a Misleading 0.00000000563422 Misleading huge financial burden for the citizens of California to bear. It s time for California, as well as the rest of the nation overall: 0.00000000561622 for that matter, to end marijuana prohibition once and for all. We cannot afford not to. Featured image from David **Degree Ranking** Degree 0.00000000523715 Return choose: 0.00000000477161 An Open Disinformation Query enormous: 0.00000000423931 Front-end Back-end **User Interface of DISCO** System Architecture of DISCO



Resources

- Online Repository: https://github.com/DongqiFu/DISCO
 - Code: Python
 - **Pretrained Language Model**
 - DISCO Demo
 - User Interface
 - Datasets: Real-World Fake News and Real News [1]
 - 21,418 real news items
 - 23,538 false news items
 - Data format: (title, text, subject, date, label)
 - Video: Use Case of DISCO





31st ACM International Conference on Information and Knowledge Management

October 17-21, 2022

CIKM2022

Hybrid Conference, Hosted in Atlanta, Georgia, USA

Thanks!



Dongqi Fu (UIUC)



Yikun Ban (UIUC)



Hanghang Tong
(UIUC)



Ross Maciejewski (ASU)



Jingrui He (UIUC)

Presenter: Dongqi Fu

Email: dongqif2@illinois.edu







