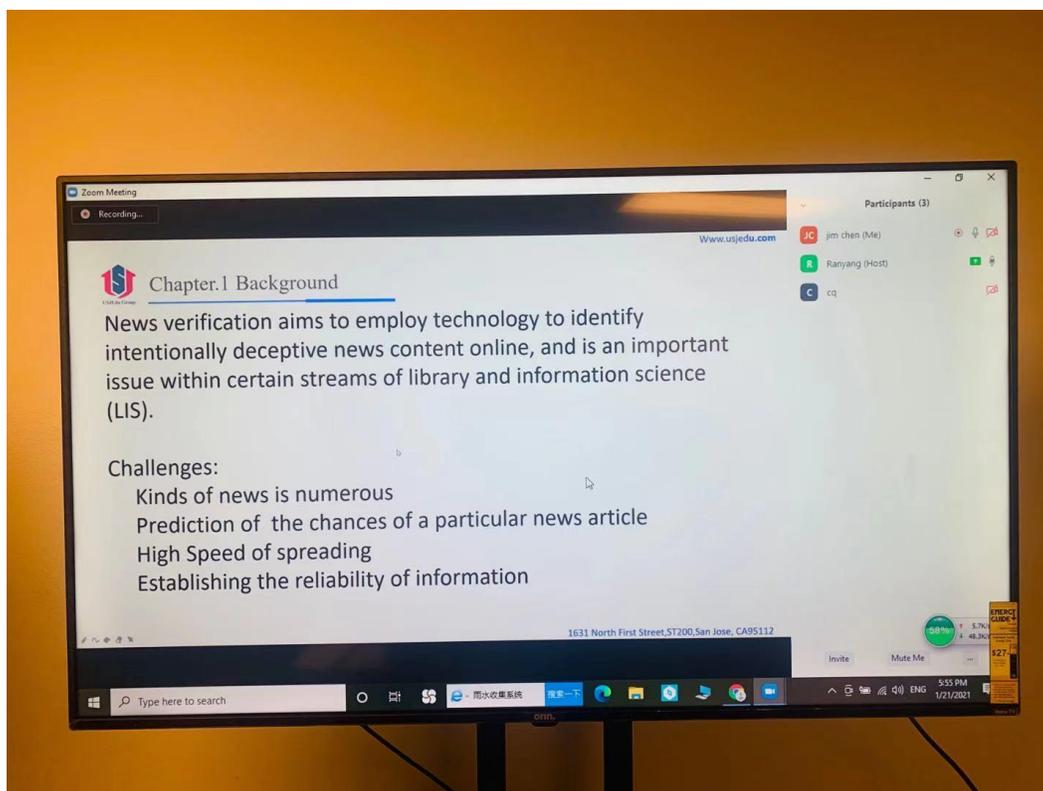


Lecture on "How to quickly and effectively identify fake news"

At 6pm to 7:30pm on January 21, 2021, the lecture of "How to quickly and effectively identify fake news" of School of Silicon Valley Artisan of USJ broadcasted live online globally. News verification aims to use technology to identify deliberately deceptive news content online and is an important issue in certain areas of library and information science (LIS). The instructor showed everyone how to quickly and effectively identify fake news. The following are the key points summarized by Ranyang, the instructor of this lecture, all are dry goods!



Recording

Chapter.2.2 Algorithm

Algorithm 1: Data Representation

"bag of words" approach:
Each word as a single, equally significant unit.

Individual word or "n-grams" (multiword) frequencies are aggregated and analyzed to reveal cues of deception.

1631 North First Street, ST200, San Jose, CA 95112

www.usjedu.com

327

Recording

Remaining Meeting Time: 00:41 | Upgrade to Pro

Chapter.2.2 Algorithm

Linked data

The inherently structured data network of entities is connected through a predicate relationship.

This approach leverages an existing body of collective human knowledge to assess the truth of new statements. The method depends on querying existing knowledge networks, or publicly available structured data, such as DBpedia ontology, or the Google Relation Extraction Corpus (GREC).

1631 North First Street, ST200, San Jose, CA 95112

www.usjedu.com

327

I. The review of lecture of “How to quickly and effectively identify fake news”

① What are the challenges facing fake news verification?

1. A wide variety of news
2. How to predict whether a specific news article is true or false
3. The rapid spread of fake news
4. How to establish a reliable information transmission channel

② 5 mainstream news recognition algorithms

1. Word Split

Split the sentence into individual words or multi-word phrases. Each word or phrase is an equal unit.

After splitting, the frequency of use of individual words or "n-grams" (multi-word phrases) is aggregated and analyzed to reveal clues of deception.

Further mark words as corresponding vocabulary cues, such as part of speech or "shallow syntax" (Hancock & Markowitz, 2013), emotional dimension (Vrij, 2006) or location-based words (Hancock et al., 2013). A method of providing frequency sets to reveal linguistic clues of deception.

2. Probability Context Free Grammars (PCFG)

The analysis of word usage is often insufficient to predict deception.

Sentences are converted into a set (parse tree) to describe the grammatical structure according to the prescribed rewriting rules.

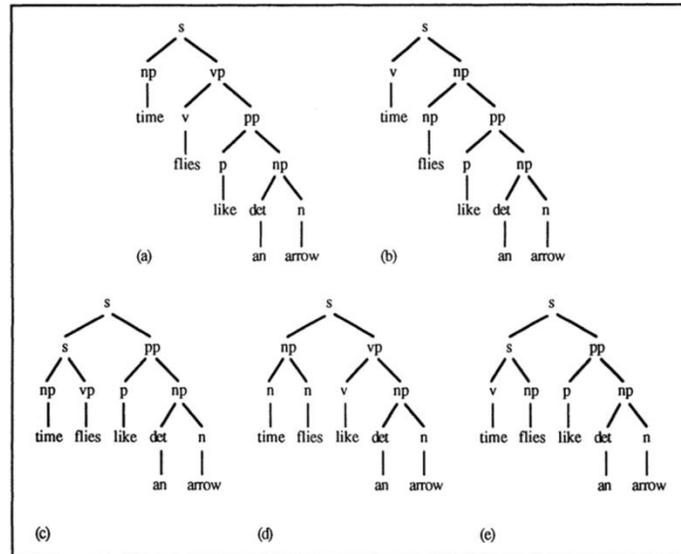


Figure 1.1. The five possible parses of the sentence "Time flies like an arrow"

③ Semantic analysis

This method expands the multi-word phrase split grammar model by merging the configuration file compatibility function, showing that this addition can significantly improve the classification performance.

④ Rhetorical structure and discourse analysis

The description of discourse can be accomplished through the "Rhetorical Structure Theory" (RST) analysis framework, which determines instances of rhetorical relationships between language elements.

The collection of word and category frequency is useful for subsequent automatic numerical analysis. A common use is to train "classifiers" in support vector machines (SVM). Using different clustering methods and distance functions between data points will affect the accuracy of SVM.

⑤ Network layer algorithm

Social network behavior: Through popular technologies such as microblogging, news dissemination in the form of current affairs news has inspired methods to determine the difference between fake and real content.

Central Resonance Analysis (CRA) is a mode of network-based text analysis, which expresses the content of a large amount of text by identifying the most important words that link other words in the network.

II. Lecture Content

1. How to break news into effective features
2. How to classify features
3. How to quickly analyze the features of many categories
4. How to prevent semantic leakage
5. Application scenarios and directions of fake news prediction

III . Lecturer

School of Silicon Valley Artisan of USJ Instructor: Ranyang



(George Washington University in D.C., MS degree in Computer Engineering

Familiar with python, working experience in Foxconn big data analysis group)

IV. Organizer

USJ SVA



Scan QR code to follow us

USJEdu Group