Introduction

Labeling fake news

- Creations of such data is expensive and time-consuming
- Accurate labels can only be obtained when the annotators have sufficient knowledge about the events.
- Dynamic nature of news articles leads to decaying quality of existing labeled samples.
  - Some of these samples may become outdated quickly and can't represent the news articles on newly emerged events.
  - Annotators have to continuously label newly emerging news articles, which is infeasible
    - It's essential to tackle the challenge of labeling fake news

Introduction

Leverage the feedback provided by users who read the news

- A news article published on a WeChat official account, a user who reads the article can report whether this news is fake or not with a brief explanation.
- Such reports from users can be regarded as “weak” annotation for the task of fake news detection
  - The large collection of user reports can help alleviate the label shortage problem in fake news detection
  - These weak annotated samples are unavoidably compared with expert-labeled samples
    - Users may report real news as fake or the reasons they provide may not be meaningful
    - Transform weak annotation to labeled samples in the training set and select high-quality samples is the major issue need to solve