

Proposal

1. The problem definition

隨著網際網路及社群網路的興起，傳統要從報章雜誌上才能獲得的新聞資訊也已經演變成隨處可得，但這也造成了任何人都可以發佈一個信息當作新聞，進而導致網際網路上充斥著各式各樣的假新聞，因此，許多人開始思考，資料科學家們可以利用哪些資訊來檢測這則新聞是否是假新聞，藉此來防止假新聞四處傳播，影響社會大眾。

2. Input of the model

新聞標題、內容、來源網址、使用語言、作者、來源國家、回復次數、臉書讚數。

3. Output of the model

Bad News:

- Fake News
- Satire
- Extreme Bias
- Conspiracy Theory
- Rumor Mill
- State News
- Junk Science
- Hate Group
- Clickbait
- Proceed With Caution

Good News

4. Possible datasets

<https://www.kaggle.com/mrisdal/fake-news/data>

<https://webhose.io/datasets>

5. Evaluation metrics

Accuracy.

Decide this news is which kind of bad news or good news.

6. References

<https://www.kaggle.com/mrisdal/fake-news>