

# Experiments

## Dataset

		# News	# Report	# Avg. Reports/News
Unlabeled	-	22981	31170	1.36
Labeled Training	Fake	1220	2010	1.65
	Real	1220	1740	1.43
Labeled Testing	Fake	870	1640	1.89
	Real	870	1411	1.62

- Experiments are conducted on WeChat's Official Accounts
- In this dataset, the news are collected from WeChat's Official Accounts (2018.03-2018.10)
- Split the fake news and real news into training and testing sets according to the post timestamp, training dataset (2018.03-2018.09), testing dataset (2018.09-2018.10)
  - There is no overlapped timestamp of news between these two sets
  - This design is to evaluate the performance of fake news detection on the fresh news
- Also have an unlabeled set containing a large amount of collected news without annotation (2018.09-2018.10)
- Note that the headlines can be seen as the summary of the news content. In the manual annotation process, experts only look at headlines to conduct labeling. Thus, in this paper, use headlines as the input data.

# Experiments

## Dataset examples

	Ofiicial Account Name	Title	News Url	Image Url	Report Content	label
0	私家车第一广播	国务院宣布：生孩子有补助了！明年1月起实施，浙江属于这档！	http://mp.weixin.qq.com/s?__biz=MTA1NTc0MjE0MA...	http://mmbiz.qpic.cn/mmbiz_jpg/j27ttKHs7TIFAL5...	[国务院没有发布过类似信息]	0
1	杭州交通918	4个年轻帅小伙突然人没了，身亡真相惊呆所有人! 太可惜了	http://mp.weixin.qq.com/s?__biz=MTA5Mzc3MDQyMA...	http://mmbiz.qpic.cn/mmbiz_jpg/0y9ibmULDTbDuCt...	[ ? ? ? ? ]	0
2	腾讯娱乐	迪丽热巴时装周走秀气场一米八，病态妆容也挡不住她的高级感	http://mp.weixin.qq.com/s?__biz=MTA5NTIzNDE2MQ...	http://mmbiz.qpic.cn/mmbiz_jpg/9Ju9PZ1NxhfdGHM...	[那个泰国人不是模特]	0
3	腾讯娱乐	李晨北京四合院内景曝光，还和妈妈一起吃饺子画面hin温馨	http://mp.weixin.qq.com/s?__biz=MTA5NTIzNDE2MQ...	http://mmbiz.qpic.cn/mmbiz_jpg/9Ju9PZ1Nxhd8SmK...	[造谣生事]	0
4	央视新闻	唾液测天赋、饭后剧烈运动得阑尾炎...8月“科学”流言 你中招了吗？	http://mp.weixin.qq.com/s?__biz=MTI0MDU3NDYwMQ...	http://mmbiz.qpic.cn/mmbiz_jpg/oq1PymRI9D7ZOQU...	[唾液基因检测的确可以找出孩子的优势潜能，明确孩子的培养方向，科学正确引导孩子发展长处，助孩...	0