

# 某社交平台谣言数据集说明

## 一、数据集内容

此数据集涵盖了某社交平台公共卫生相关谣言及其特征，针对特征，主要包含了用户特征以及文章传播特征，主要的字段有：发文数、关注量、粉丝量、点赞量、评论数、谣言率、负面率等。通过分析用户信息，可以更好地判断出用户评论的真实性，同时筛选出发布谣言多的用户，从而维护社会秩序和公共安全。

## 二、基本字段介绍

字段	描述
postid	文章的唯一标识符
发文数	发文人的发文数量
关注数	发文人的关注数量
粉丝数	发文人的粉丝数量
introduction	发文人是否有简介（0：否，1：是）
是否认证	发文人是否有认证（0：否，1：是）
点赞数	文章的点赞数
评论数	文章的评论数
是否包含特殊符号	文章是否有特殊符号（0：否，1：是）
positive_prob	文章内容属于积极类别的概率，范围 0-1
sentiment	文章的情感极性分类结果，0:负向，1:中性，2:正向
confidence	文章的情感极性分类的置信度，范围 0-1
negative_prob	文章的内容属于消极类别的概率，范围 0-1
abnormal	文章是否为谣言，0：否，1：是