**基于**BERT**神经网络模型的智能网络谣言监测系统**  
分析报告

**INRMS Analysis Report**

数据集：我是数据集名; 生成日期：我是日期; 数据总量：我是数据总量

# 数据集基本情况

数据集的数据量为：我是数据总量 其中，在日期上的分布如图1所示。

.图1.

此图描述了样本数据集的分布情况，可以较为直观地观察该话题的热度情况。柱状图越高，表明在该天内网络条文的数目越多，说明网民的讨论积极性在该天较高。

# 模型训练

本次任务 [进行了] [未进行] 模型训练， 经过训练，数据真实性的分布如图2所示

.图2.

如果进行训练，BERT模型得到的结果将展示谣言可能的分布状态。注意，应当有一个结合具体实际情况的参数才能得到较为准确的结果。所有参数在系统界面内都可进行调整。训练结果保存在工作目录下test\_results.tsv文件内。该图片说明了此数据集中，谣言（或不实信息）所占的比例。

# 数据分析

综合情感语料库分析，在阈值为 我是阈值 的情况下，态度状况随日期的分布如图3所示：

.图3.

语言激烈程度状况随日期的分布如图4所示：

.图4.

态度以及语言激烈程度说明了在数据集网络条文中对该话题积极或消极的态度，往往这个态度会随着时间进行而有所变化。舆情的产生，不实信息的流传往往也会伴随着态度和激烈程度的波动变化。

他们在总数据集中的占比如图5、图6所示

.图5.

.图6.

一般情况下，积极态度和非激烈的语言应当占数据集的大多数。如果有所异常，该数据集的来源网站或论坛等或应当引起重视，慎重研判。

# 预警信息

将以上数据进行数据归一化操作，综合分布如图7所示

.图7.

数据所见：

我是数据所见

预警提示：

我是预警提示

# 辅助判断信息

通过对次任务数据集中的语言词汇进行统计分析，词频以**词云的形式**展现如图8所示:

.图8.

词云展现了网络条文中的网民所讨论的具体主题。频率较高的主题词汇将会被以较大的字体展示。在参考分析结果的同时，请注意同时适度把握具体的舆论情况。

智能谣言监测系统生成报告

*No Bug Workgroup Version* 0.1.1

我是日期