

预测灾情推文的真实性

组员

| | | |
|---------------|------------|------------|
| 名字：梁景辉 | 冯杰晖 | 王则刚 |
| 学号：2006500016 | 2006500011 | 2006500015 |

kaggle 链接

[Natural Language Processing with Disaster Tweets | Kaggle](#)

背景

社交媒体平台如 Twitter 等已经成为广泛传播信息的主要渠道之一。这些平台可以让用户在实时，快速且广泛的情况下分享信息，特别是在自然灾害和突发事件发生时。然而，由于信息来源的不可靠性，社交媒体上的许多消息可能是错误的或具有误导性的。因此，一个有趣的问题就是如何确定哪些 Twitter 信息是有关实际灾难事件的。

选题意义

这个问题具有重要的现实意义。当灾难事件发生时，能够及时准确地识别哪些信息是可信，哪些是虚假的，可以帮助政府和救援机构更好地了解灾害范围和灾难的严重程度。同样，对于需要离开自己所在地区或国家的人来说，可靠的信息可以有效的帮助他们决定何时离开以及应该去哪里。

想要解决的问题

社交媒体中的信息往往是非正式和不可靠的。这使得区分真实和虚假的信息变得困难，尤其是在灾难事件发生时。因此，解决这个问题的核心是如何判断哪些信息是涉及实际灾难事件的。

研究内容

数据集：

我们将使用 Kaggle 上的 Twitter 数据集，其中包含有标记的 Twitter 信息，以何种方式判断它们是否与真实灾难事件相关。

算法：

我们计划使用 transformer 对 Twitter 文本进行分类并确定 Twitter 信息是否与灾难事件相关。

首先，收集 Twitter 文本数据，对文本进行预处理（如数据清理等）以进行分类。这些文本可以是关于灾难事件的推文，也可以是与灾难无关的推文，还需要准备好正负样本进行训练和测试模型。

其次，使用 transformer 模型（预期将会对注意力机制进行一定修改）对收集的数据进行分类。此外，还需要确定模型的结构、超参数和训练方法等。

最后，使用训练好的模型对新的 Twitter 信息进行分类，以确定是否与灾难事件相关。在这个过程中，可能需要处理文本数据的预处理和后处理，如分词、去噪声等。同时，还需要对模型的准确率、召回率等进行评估，以确保模型的性能。

预期目标：

预期目标是准确地识别与灾难事件有关的 Twitter 信息，从而帮助准确，及时地了解灾害范围和程度，帮助人们做出适当的决策。同时，我们也希望将所开发的模型扩展到其他自然语言处理任务中。