**Questions To Stimulate Class Discussion**

1. **How might the performance of the models change if the dataset is significantly larger or comes from a more diverse range of disaster scenarios?**

The performance of models like CNNs and LSTMs generally improves with larger, more diverse datasets. More data offers a wider range of examples and contexts, helping these models generalize better to new, unseen scenarios. This is especially important in disaster scenarios, where variations in language and context are significant. Training with diverse datasets covering different types of disasters, regions, and languages enhances model robustness, reduces overfitting, and increases accuracy, ensuring the model performs well across various situations.

1. **How can we effectively integrate human judgment with automated systems in the process of classifying disaster-related tweets to enhance both accuracy and trust in the system?**

First, include a system where humans check the machine's decisions whenever there's uncertainty or in critical situations. This helps improve accuracy and also provides feedback to enhance the system. Secondly, set up a group of experts who can regularly update the rules the machine uses to classify tweets, keeping it up-to-date with new types of disasters and trends. It’s also important to have a process in place that quickly brings certain high-risk situations to human attention. Additionally, providing regular training for the people reviewing the machine’s work helps them better understand and make use of the AI’s suggestions while being aware of the context. Lastly, being clear about how much the machine is involved and how humans are included in the process builds trust among users and stakeholders. These steps help ensure that the strengths of both humans and machines are used together, making the system more reliable and trusted.