Notes from Duke DS+ session on NLP for covid literatures

Document classification:

* Sentiment analysis: binary classification problem (binary classification is easier so generally start here)
* Data screening: label as relevant or not-relevant, useful as a preprocessing step for other things