

WHAT IS NLP?



The Perfect Recipe

“Natural-language processing (NLP) is an area of computer science and **artificial intelligence** concerned with the interactions between **computers** and **human (natural) languages**, in particular how to program computers to fruitfully **process** large amounts of natural language data.”



Fake News Challenge - 1



Classifying Articles

1. **Agree** with the article?
2. **Disagree** with the article?
3. Discuss the same topic as the headline, without the article taking a stance? (**related**)
4. Discuss a different topic than the headline (so, **unrelated**)?



What did we do last week?

- Review rule-based classifications?
- Evaluation metrics
 - Accuracy
 - Bad for uneven distributions of labels
 - True positive, false positive, true negative, false negative

$$\text{Precision} = \frac{tp}{tp + fp}$$

⇒ make sure that the ones classified as positive are actually positive

$$\text{Recall} = \frac{tp}{tp + fn}$$

⇒ make sure to catch the ones that are actually positive



Harmonic Mean

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

