**Bonus: Behind the Scenes Sneak Peak**

**We identified two broad principals of calculating the authenticity of a content piece.**

**Principal 1: Tally with renowned news sources.**

Working: Viral extent of a news can be taken from trending tags on social media and other news

sharing apps like newshound, InShorts, etc. Viral extent maybe taken into account for

determining whether news is fake or not. Tags and key words from the headline and content of

the news can be used in a search algorithm that checks whether an article with same keywords

is present on the renowned news sources or not. Accordingly a news can be declared fake or

genuine. For determining a renowned source basic criteria such as year of establishment, user

reviews, etcetera are taken into account.

**Pros:** No human interaction.

**Cons:** Might show real news fake if no occurrences found on renowned sources. Can be overcome

by little human input.

**Principle 2: Human resource**

Can be implemented in two ways:

1. Dedicated experts

2. Crowdsourcing

**1. Dedicated experts:**Dedicated experts can review after an algorithm does basic filtration.

**Pros:** High accuracy

**Cons:** Slow process, expensive deal

**2. Crowdsourcing:** Volunteers who act as moderators can flag an article fake or genuine and according to the

reports fake news can be segregated from the genuine one. There could be a reward, rating, and

recognition system that motivates users to flag the fake news.

**Pros:** Can be accurate to a large extent

**Cons:** Cannot trust all volunteers but if sample size is huge, average reviews can result in

accurate flagging of the news.

**DETERMINATION OF PERCENTAGE OF AUTHENTICITY**

Over time the algorithm can learn which keywords are highly likely to be used in a fake article.

On this basis, it will increase the confidence of the percentage determined by the algorithm and

consequently decrease human intervention. The weighted mean of all percentages will give the

final percentage.

Confidence of P1 (by computer) = x

Confidence of P2 (by computer) = y

Confidence of P3 (by computer) = z

Where P(i) is the percentage

Final Percentage P = P1x + P2y + P3z

As the computer keeps learning to identify keywords, genres and trends of highly faked news, it

will ask for human intervention for fewer links.

P1 is a function of number of keywords that match patterns of highly faked news, difference

from original source, reliability of original source, genre and trend.

P2 = no. of moderators who marked authentic/ total no. of moderators asked

P3 = no. of people who marked authentic/ total no. of people asked

One such algorithm can be made that understands the context and accordingly matches latest

news from renowned sources and finally determines the percentage of authenticity. For

example, fake news is that a celebrity died in a car accident. The algorithm understands the

context as tragedy and car accident and look for similar occurrences. Say no such keyword was

found that suggests that the celebrity actually died in that. Therefore, taking all the information

into account a percentage of authenticity can be determined.