A chat app with explainable fake news detection

10.10.2021

—

Amit Kumar Gupta
universität siegen
Adolf-Reichweine Str. 2
Siegen, 57072

Overview

According to statistics Germany has 71.1% smartphone penetration in its population. Each smartphone has a chatting app installed onto it. Having a smartphone on every hand is a positive event but there are certain events which can trigger a massive spread. The spread of fake news is something which is bothering the governments a lot. It is extremely hard to detect and stop the spread of fake news into the masses. Especially in the age of 5G, where we have extremely fast and cheap internet access.

To automate the process of detecting fake news and stop the spreading of it is a tricky and complex task to be achieved. The main challenge would be to convince the end user that the message received is fake news. Otherwise deleting the message without the consent of the user would turn out to be a drastic step as it would be against the right to express. So it is very important for the developer to develop a detection algorithm with a hint of explanation of the detection. To make the user more accountable for sharing the fake news would be to make him/her know about the fact that the message which he/she has received and going to send to other users is fake news.

To achieve this task the first step is to classify the message received as news or NoNews. After the classification if the message is classified as news then the message will be passed to the next algorithm where the message will be classified as fake or real. So to classify any particular message into fake or real, the message has to go through 2 separate classification algorithms.

Now the next challenge would be to make the smartphone user convinced over the classification decision made by the algorithms. This is the major task of this project. Explanation of the model output has to be made and that explanation should be understandable to the user. Since users come from different backgrounds they might not have the expertise to understand the technicalities of working of the algorithm. The initial step to this would be to run classification algorithms before the other two mentioned above. In this case EmotionAI algorithm would run and try to classify the sentiments of the messages received.

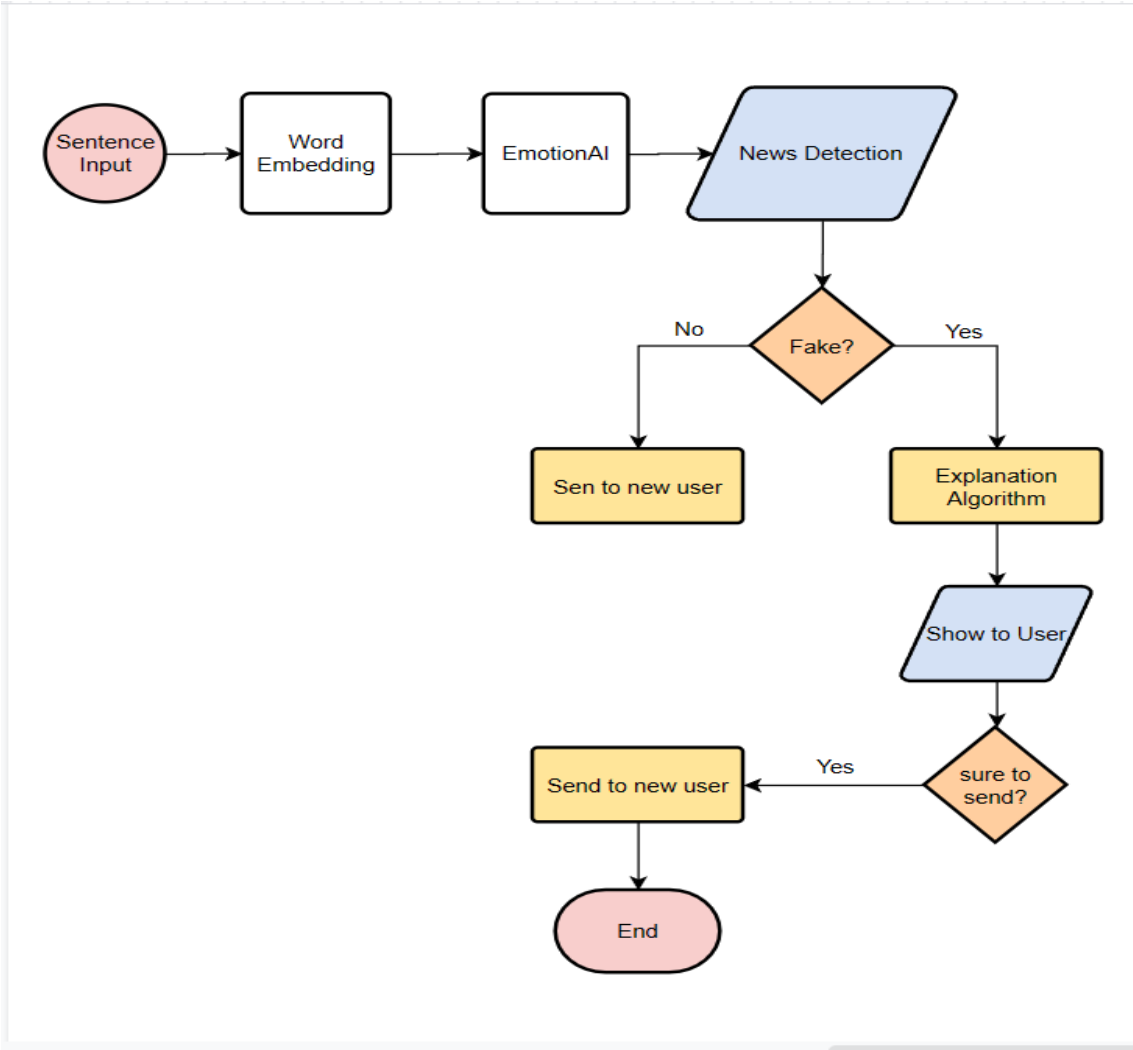

Fig.1 : Fake News Algorithm

Goals

1. End goal is to make a web/mobile app where a person can chat and share news articles with fake news filtering.

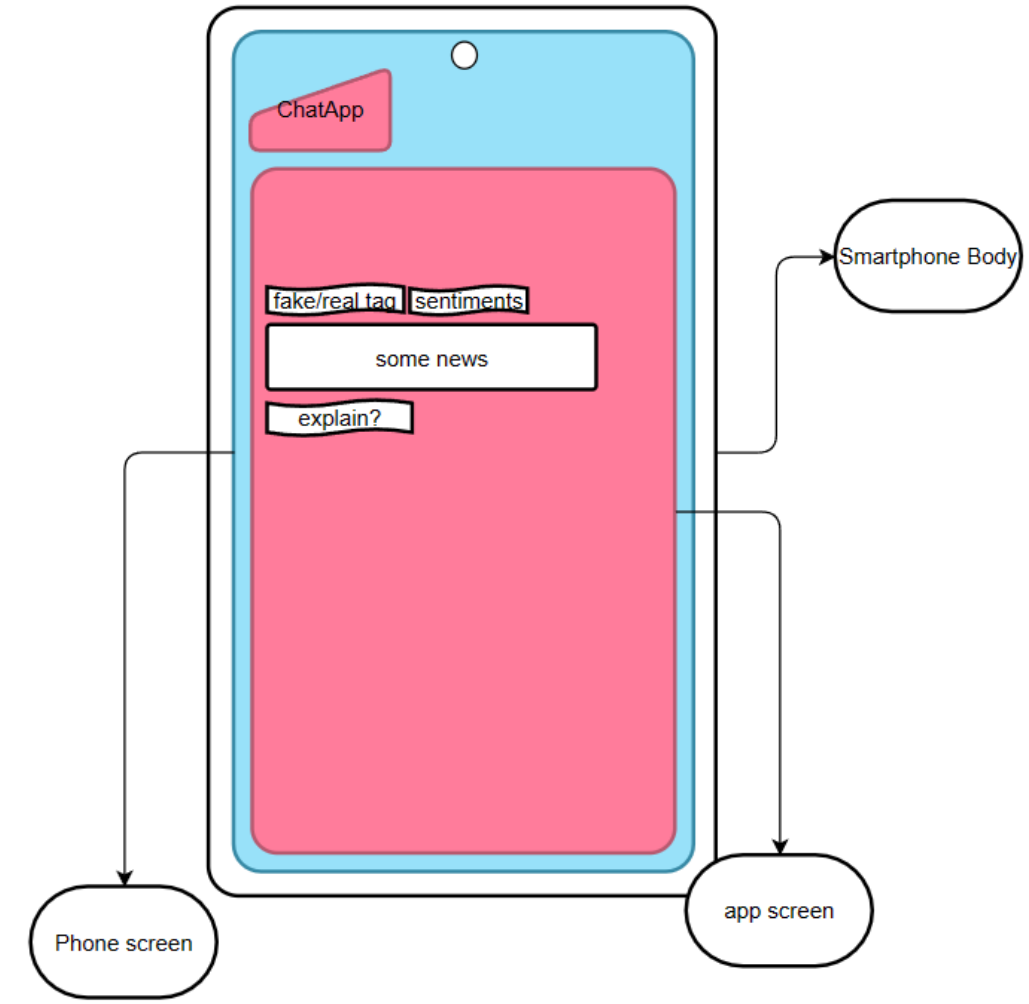

Fig.2 : Initial chat app outline

2. Explain the fake news classification using methods such as Knowledge graph.