

Case Study Name: Spread of fake news in social media video content creation in India

Study By: Gourav Baloria

Date: March-2024

Objective: Our objective is to analyze the spread of misinformation using social media channels through videos. To prevent it, develop a policy recommendation for online video platforms and government to control the fake news.

Why do we need?

- With massive growth of apps like Whatsapp, Facebook, tiktok, YouTube, Instagram etc in short time span, content now reaches wider demography. It makes crucial to study fake news separation pattern.
- It's concise and visual nature, help it spread rapidly to wider audience. This enables the risk of online hate spreading, communal violence, economical losses etc.
- It's visual nature and short duration lend credibility initially before proper verification. Therefor It is difficult to do fact check before it reaches to wider audiences.
- To prevent take political advantage by the parties against their opponent by establishing their IT operations for spread such content which help to gain funding, win election etc.

What is the core need?

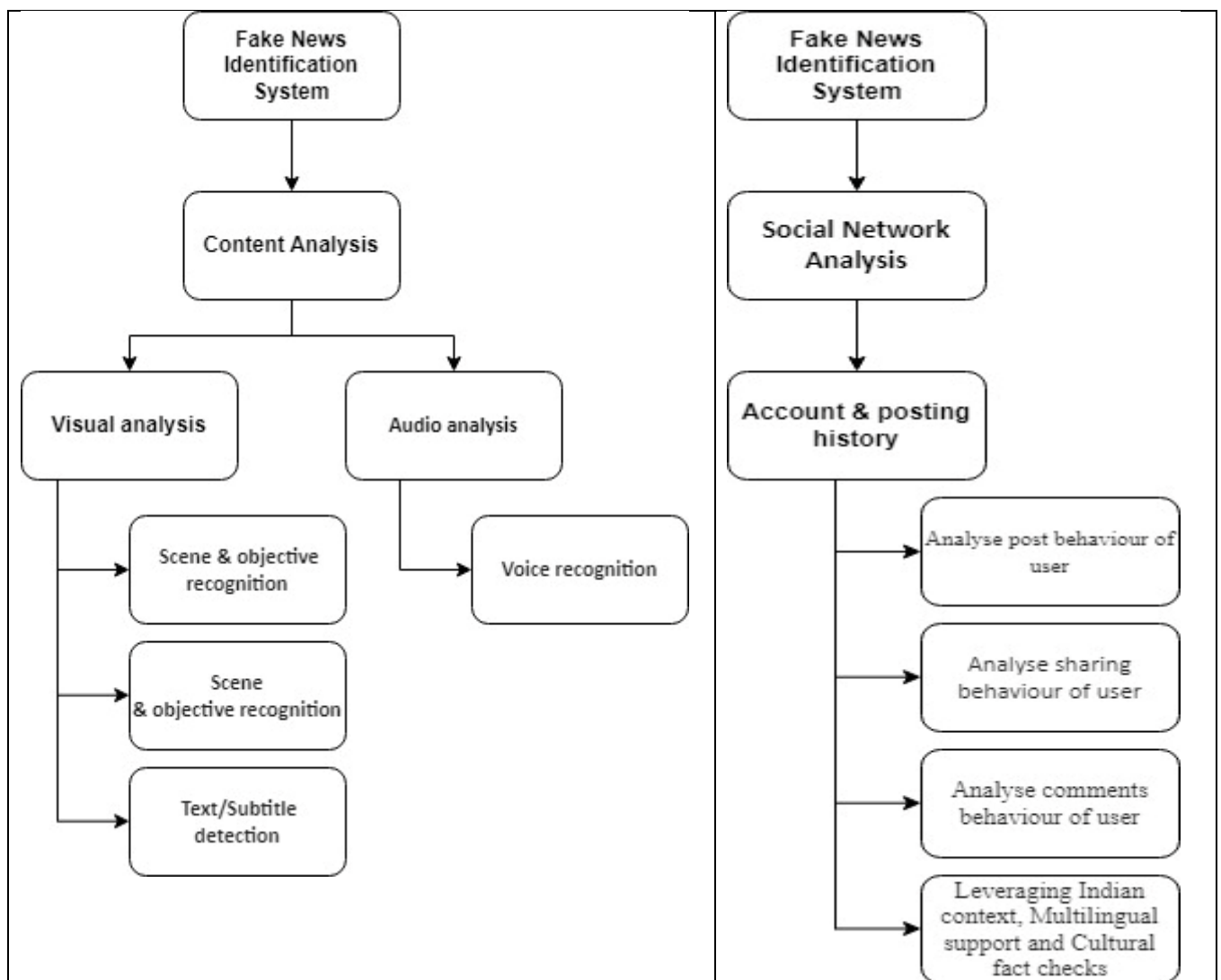
- Over 50% of Indians are consuming internet every day. It became essential to draw a framework for the companies and consumers to prevent the fake news. [Link](#)
- Many reports based on internet, in 2019 general election approximately 30% content shared on social media was found fake or misleading. [Link](#)
- Developing an effective, evidence based a policy framework for digital platform to control the spread of misinformation through better detection, initiative-taking prevention of the fake content.
- YouTube stated they removed over 600000 videos and terminated 1.5 million channels for policy violation around misinformation in the first quarter of 2012 around globe. Meter specific to India is unclear. [Link](#)
- India was documented having one of the highest WhatsApp misinformation-related lynching cases in 2018, prompting policy changes. [Link](#)
- Around public events of many popular leaders and events, fake news related to bomb blast, murder etc.s are spread by people at large scale via short video content for gaining attention to their channels/ pages.

So, in summary, existing matrix struggling with overall misinformation, especially around key events like election, large rallies, public events etc.

Challenges in the existing matrix

- Lack of educational initiative from these video platform owners to educate their user base about the policies. Awareness and resources around misinformation prevention policies for creators is minimal.
- Lack of platform accountability, video platform does not own responsibility for misinformation spreading from there platform.
- Fact checks for every video is impossible given the volume and speed of creation. Automated mechanisms lag in Indian language support compared to platform usage growth.
- Minimal identity verification requirements for users generating content enables them to create misleading content through anonymity.

Solution and its scope



Solution Brainstorming:

- **Content Analysis**
 - **Visual analysis**
 - **Deepfakes detection:**
 - As deepfakes become increasingly sophisticated, developing robust detection systems is crucial. Training these AI models requires significant data, funding, and expertise. Collaboration between governments and companies with AI specialists is essential for this endeavour. AI can use features of people already available on social media like height, age, weight, skin colour, ethnicity etc. to identify inconsistencies or manipulations in video frames that might indicate deepfakes or edits.
 - **Scene & objective recognition**
 - Many time people share videos of one particular events from other place by associating with another, if it has any political, religion or any other kind of co-relation. Therefor a system can get developed which identify time/date, location, events of the fake videos based of geographic location tags present in videos, background objects in videos, identify objects that wouldn't be present in the claimed scenario and look for objects that should be there but are missing. Inconsistencies with the claimed setting or impossible situations could be red flags.
 - **Text/Subtitle detection**
 - Detect and analyse text overlays in videos with audio, especially those appearing/disappearing rapidly, which are common tactics in fake news. Identifying manipulated audio like voice cloning or dubbed speeches can further expose fake content.
 - **Audio analysis**
 - **Voice recognition**
 - Speaker Identification of the speaker(s) in the video and cross-reference with known voices of public figures. Inconsistencies or mismatch could be a sign of manipulation. Enhance the system to identify previously used audio clips (speeches, sound effects etc.) that might be spliced into new videos to create a false narrative.
- **Benefits of content analysis**
 - Protecting Reputations
 - Media and journalism integrity
 - Identification and block the fake content at early stage.
- **Challenges in content analysis**
 - Collect a massive dataset of real and fake news videos. This is crucial for training the AI models to distinguish between genuine and manipulated footage.
 - Tactics challenges, people can use real locations to make fake videos
 - To train and develop such AI model needed significant data, funding, and technology expertise. It also required massive collaboration between government, fact check organisations, media and video platform owning companies.

- Keeping up with deepfakes with evolving technology
- To train and develop such AI model needed significant data, funding, and expertise

- **Social Network Analysis**
 - **Account & posting history**
 - **Analyse post behaviour of user**
 - Analyze the accounts and keep looking for patterns of inauthentic behavior like sudden bursts of activity, unusual posting times, or a history of sharing fake news etc.
 - **Analyse sharing behaviour of user**
 - Keep track of unknown or unreliable sources, fringe websites so when user share any information from such sources it can get easily identify. Also, make a previous sharing history report for channels/pages etc. based on sharing from unreliable and fringe sources.
 - **Analyse comments behaviour of user**
 - Identify users who insult others, use aggressive words to evoke strong emotions. User posting same comment on number of posts etc.
 - **Leveraging Indian context, Multilingual support and Cultural fact checks**
 - Develop AI models that can detect fake news specific to Indian politics, social issues, and religious sensitivities. Track the trends/events when fake news circulation is high.

- **Benefits of Social Network Analysis**
 - Easy to identify if we compare with content analysis.
 - Empowering users to identify fake content.
 - Detecting Communities which are participating in fake news circulation.
 - SNA can be used to trace the origin of fake news stories.
 - Pinpoint influential users who spread information quickly.
 - Analyzing the content being shared can offer valuable context for understanding how information flows.

- **Challenges in Social Network Analysis**
 - Obtaining accurate and complete data on social networks can be difficult.
 - People might not truthfully represent themselves online, and some connections or interactions may not be captured.
 - SNA often involves collecting personal information, raising ethical concerns about privacy and user consent.
 - SNA utilizes complex analytical techniques that require researchers to have a strong understanding of the methods and tools used to interpret results accurately.
 - Capturing this dynamic nature can be challenging for analysis.

- **Quick Solution to start damage control**
 1. People have been caught for making fake videos, make them viral etc. recently, but there are not we still need tougher laws to stop future crimes and make it clear what

kind of trouble you can get into for doing this. Therefore government should reform the Information Technology Act and Copyright Act.

2. Government should mandate to the social media companies to setup below department
 - a. Setup fact check team and publish the identified fake posts.
 - b. Provide option to mark post as “Fake” to users.
 - c. Make a policy and set a maximum number and after that ban the account who violated and shared fake posts.
 - d. User education policies and interface
 - i. **User education**
 1. Educate user for fact check for any sensitive information
 - ii. **Transparency**
 1. Highlight the post with Fake Flag, if it is fact check, reported as fake from large audience etc.
 - e. Set a matrix to give credibility score to the user who are posting content on these platforms.
- **Remember:**
 - **Privacy concerns:** Developing such system involves collecting personal information, raising ethical concerns about privacy and user consent. Researchers need to ensure they have proper authorization for data collection.
 - **Maintaining platform neutrality:** Be neutral towards everyone and don't favour anyone.