

Big Data, Artificial Intelligence and Machine Learning: A Paradigm Shift in Election Campaigns

Md. Safiullah* and Neha Parveen

Chanakya National Law University, Patna, Bihar, India

Abstract

The paradigm of the election campaign has shifted around the world. Artificial Intelligence and Machine-Learning are now dominant over traditional ways of the election campaign. Artificial intelligence and machine learning analyzed the big data to understand voters' unique psychographic and behavioral profiles. They persuaded voters to vote for a particular politician/Political party and create a bias against political opponents during the election. Deepfakes, Automated social media bots are among the famous developments in technology that are now highly used to manipulate and alter voters' opinions during the election campaign. The present study examines the use of Artificial Intelligence and Machine-Learning during the election campaign. The present study is divided into three sections. The first section highlighted the status of internet penetration and social media's influence on the *Indian Lok Sabha* election. The second section emphasizes the use of Deepfake and automated social media bots during the election campaign. The last section explores the future of Artificial Intelligence and Machine-Learning in the election campaign.

Keywords: Artificial Intelligence, machine learning, Deepfake, bots, big data, election campaign

*Corresponding author: Saif.sheekh@gmail.com

11.1 Introduction

Technologies are developed to make human work easier. It is now used by organization to understand their customer. Technology has changed the election campaign and advertisement approach around the democratic countries. A combination of big data, Artificial Intelligence, and Machine Learning can provide crucial insight and uncover essential facts that can be immensely beneficial for the election campaign. Big Data helps political parties to read voters' minds and patterns of behavior. Social media and mobile applications are considered the primary source of big data.

According to the survey, there are 91.10 crore voters in India during the 2019 Indian General Election. Around 54 crores are mobile phone users having Facebook and WhatsApp accounts. Social media can influence 30% of total voters during the 2019 Indian general election [1].

In election season, big data is worth its gigabytes is gold because it allows customized campaigns by targeting individuals. Political parties in democratic countries are working with massive data sets of online behavior to understand households in a constituency. Traditional campaigning techniques exercise by politicians and political parties to achieve their objectives are consider no more relevant and valuable [2].

In recent years, technology has changed the whole approach of the election campaign. Political parties and Politicians are now more relying on new technologies. Big Data, Artificial Intelligence, and Machine Learning are now being used in analyzing psychographic and voters' behavior.

After the advent of information and technology, political campaigns have become data-driven. Now election campaign more relies on big data. Political parties across democratic countries are analyzing big data to understand the profiles of the constituency. Political parties and election strategists are more interested to know who first-time voters are. Who are floating voters, and how do they behave, demographic and socio-economic conditions of voters? Particularly in India, Political parties and political marketing consultants are more interested in understanding the caste, Religion composition in society, and voters' socio-economics conditions. These data help design election campaign strategy (What to talk with a particular cluster of voters), select candidates, and build a pro-party narrative [3].

Data related to the users' particulars are now collected with the help of technology. Software applications Google Maps, YouTube, Instagram, Facebook, Twitter, and many more applications contain tons of users' data. Advertisers have been exploring ways to utilize social media data as much

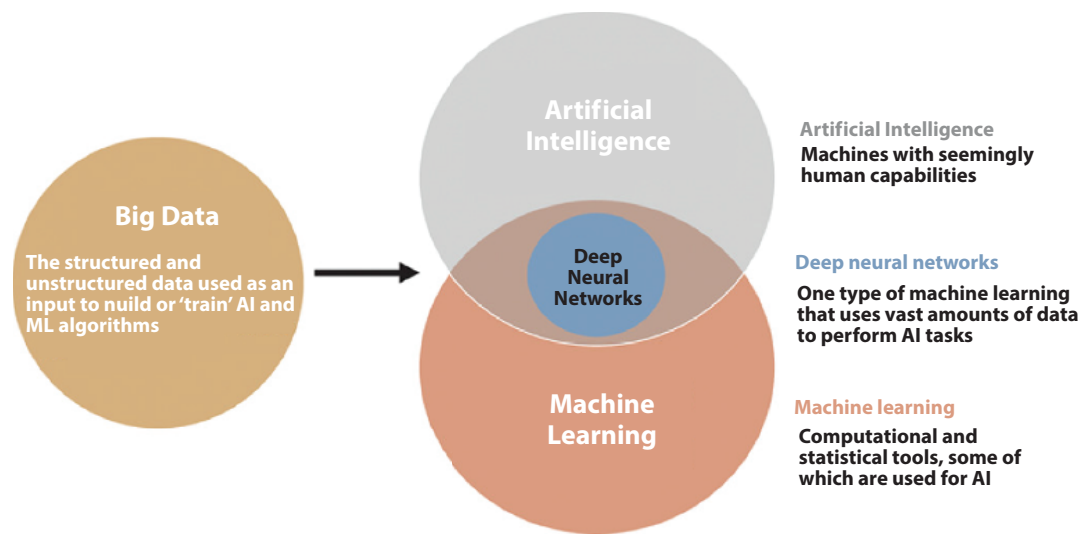


Figure 11.1 An overview of the relationship between artificial intelligence (AI), machine learning (ML), big data, and deep neural networks. Source: [5].

as possible to understand consumer behavior. Hyper-targeted advertisements are based on online and in-person activity is expected. At this point, what is new is the political parties and politicians are catching up to the advertisers [4].

Now, advertisers and political campaigners use the structure and unstructured data as an input and computational power of statistical tools to build or train machines' seemingly human capabilities. Big data, Artificial Intelligence, Machine learning, and more sophisticated political communication are now used by political parties/Political consultants to target voters more individually. The relationship between Big Data, Artificial Intelligence, and Machine Learning is depicted in Figure 11.1.

11.2 Big Data Reveals the Voters' Preference

The world is becoming a digital space. Today all works are managed, share, and stored online. Data is gathered from the device like laptop, smartphone are called big data. Big data have a particular pattern if analytical tools analyze it. These data patterns are meaningful and help in understanding the behavior and trends of individuals or groups.

Big data are now in high demand more in election campaigns to understand voters' choices and issues. Data is also used to predict election outcomes and design micro-targeting election campaigns [6]. Obama explored the potential of big data in 2012 for his US presidential election campaign.

Now, other democratic countries' political parties understand and utilize the potential of big data for the election campaign [7]. The most dominant means of collecting big data is social media and mobile applications. Figure 11.2 shows “Big data analytics examines large and different types of data to uncover hidden patterns, correlations and other insights”.

Recently, political parties have developed their own software application to connect voters' and collect big data to understand voters' perspectives, behavioral and psychographic profiles.

Citizens involved in the election are the organizer, volunteers, or supporters. Political parties demand their volunteers and supporters to download their software applications. The Apps subscribers provide their phone numbers and allow apps to access contacts numbers, photos, video, messages, GPS, and Bluetooth data. Thereby, the subscriber receives stream text, messages encouraging the subscriber to be more involved in the election campaign.

In the 2008 US presidential election, Barak Obama released election campaign apps for the iPhone. The apps access users contact list from their phone that lives in different locations. Then encourage people to call them on behalf of the political party. GPS was a new feature for the smartphone back then. GPS help to direct users to the local campaign event and campaign/political party offices [9].

All the tiny things like the user contact list, sent text messages, and received GPS location history was collected, centralized, and out of power

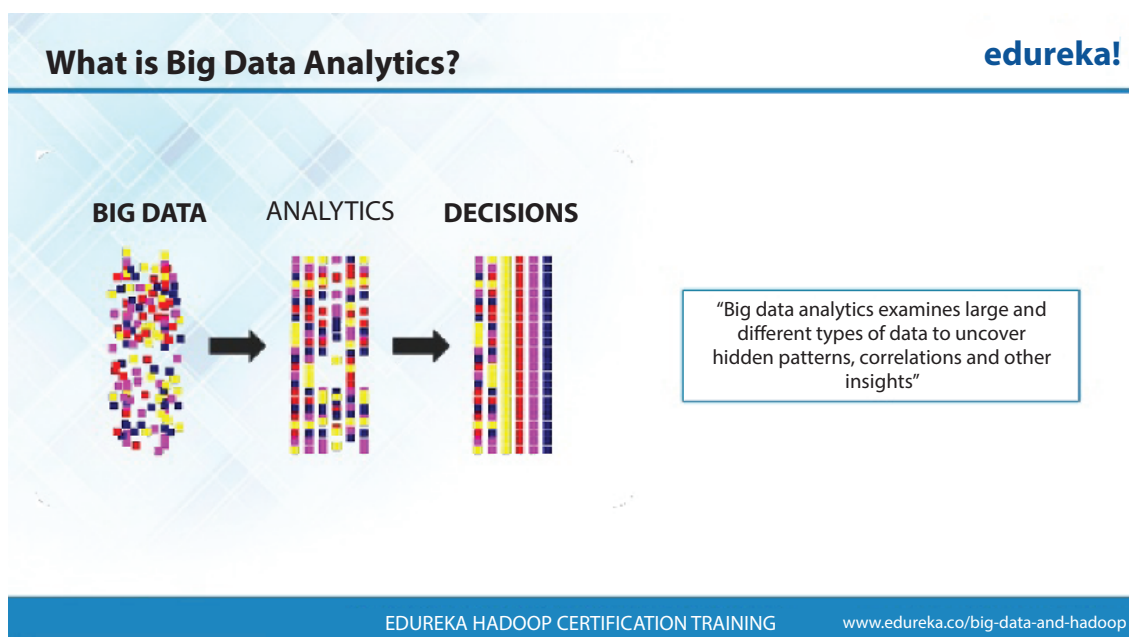


Figure 11.2 Big data analytics and visualization. Source: [8].

how the users use. This is because the user has permitted the apps to their personal information from the phone. Same we see today, it is prerequisite criteria to allow apps to access the phone memory. The smartphone had successfully replaced the door-to-door campaign to the same extent in urban America. The 2008 Obama victory in the US presidential election is considered data have driven campaign.

Obama canvassing apps give volunteers, list of neighbors to target. During the 2008 US presidential election, 1 million people gave campaign access to their Facebook data and their friends' Facebook data. Obama email lists 20 million Americans.

In the next election, Cambridge Analytica (A British political consulting firm) had obtained Facebook data on tens of millions of Americans without Facebook users' consent. Then used Facebook data to build a "psychological warfare tool," which unleashed US voters with fine-tune messages to help elect Donald Trump as president [10].

The political consulting and data analysis firm Cambridge Analytica obtained Facebook users' personal information through the third-party quiz app. Those who downloaded the app were told that the data collected would be used for academic purposes. But, these data were used to build detailed voters' profiles for both the Ted Cruz and Donald Trump presidential campaigns. Skeptical voters were then targeted with provocative political content and ads before the 2016 US presidential election. During the 2016 US presidential election, Campaign apps were getting more sophisticated. The Hillary 2016 app was heavily gamified. App users could complete each other to earn rewards [11].

Overall, the app's scandal and success in the past election have laid the foundation for robust digital election campaign strategy application.

11.2.1 Use of Software Applications in Election Campaigns

Phunware Company has developed application software for Trump 2020 election. *Phunware* Software Company is well known for its advanced location tracking capabilities Trump 2020 app is a gamified tool, news aggregator media creator, and virtual event platform with maximums approach to data collection. This application software provides facilities like money donation facilities, social media conversations, surveys, and many more. This application software is beneficial for target advertising in commercial space. Location-based services help track user activities like the stores' choices and time spent on the particular location and visualize user movement patterns in the set boundaries. This information is then triangulated with other data and developed personalized advertising [12].

MIT Technology Review, which evaluated both application software, said Trump's "Data collection is perhaps the most powerful thing the Trump 2020 app does." A phone number for a verification code and the user's full name, email address, and zip code are required to sign up in the software application. Users are also encouraged to share the app with their existing contacts. This campaign strategy helped reach around 40 to 50 million citizens expected to vote for Trump's reelection. This campaign strategy intended to collect every single one of these voters' cell phone numbers. This strategy means the app also makes extensive permission requests, asking for access to location data, phone identity, and control over the handset's Bluetooth function [13].

11.2.1.1 Team Joe App

For Joe Biden, his campaign has launched the "Team Joe" app. According to the Biden campaign, the app's primary purpose is about community and voters sharing their experience and their support for the candidate with their network. This type of relational organizing leverages volunteers' influence and personal contacts, giving the campaign access to potentially like-minded individuals.

"The Team Joe app is an organizing tool that allows users to text their friends in support of Joe and get updates from the campaign," according to the Team Joe app website. "App will let the user know which of user friends and family members we're hoping to talk. Users can then text them directly to share campaign updates, ask questions, and gather their thoughts on the 2020 race on behalf of our campaign" [13].

11.2.1.2 Trump 2020

In the app store, Trump 2020 app is categorized as a news app and rated 17plus revision, including mild profanity or crude humor, mature/suggestive themes, medical/ treatment, and information. According to Apptopia (a web-based market data and intelligence solution for mobile application developers), "It provides several data points such as revenue, app store downloads, and overall performance with statistics for every mobile app." This app also offers intelligence capabilities to identify emerging trends and consumer interests.), it has over 2.6m downloads in early October 2020, over 15 times of Biden app (Vote Joe). Biden app is categorized as a social networking app and appropriate for ages four and above. Figure 11.3 shows the app store preview of the Trump 2020 app and the Biden app [14].

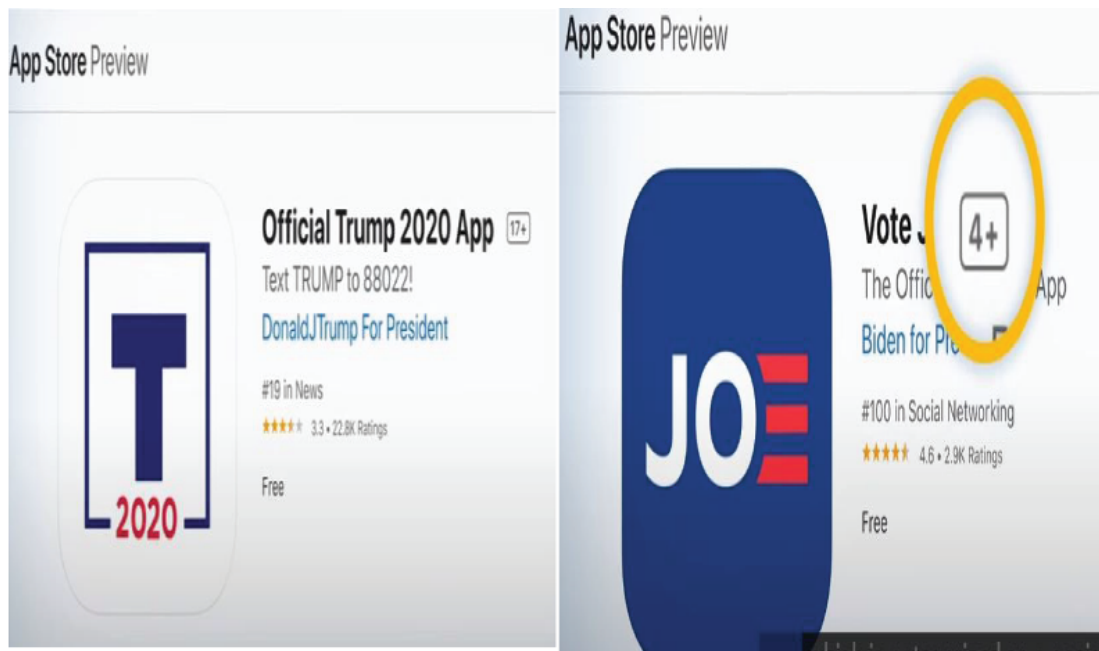


Figure 11.3 App Store Preview. Source: [14].

The difference between Trump and Biden is based on its approach. Trump’s app seemingly targeted all its supporters, while Biden app is designed for one purpose, “relational organizing.” That volunteers and organizers strategically reach out to friends and family to vote for Biden.

It’s a coin with two sides. The best way it changes someone’s mind about something they believe in politically is to talk to them and know them. But the best way to manipulate someone is to get some care about and share information with them. This strategy depends on the message being passed.

Never “relational organization” is a campaign with high volume semi-automated average tools. If changes, scale changes automation. So text friend required a friend to be registered on the app. This function is built on an automated system that turns it into a new form of propaganda.

11.2.1.3 *Modi App*

March 2015 Modi app has been downloaded 10 million times. Default permission gave total access to all the data on the user’s phone. The app is often used to disseminate news and propaganda to supporters. The Trump and Modi apps appealed the media ecosystem. One of the objectives of the app is to prove mainstream is corrupt. The user gets attention-grabbing articles from the campaign media team. Modi app takes 22 permissions from the user, including the ability to access the user’s contacts, gallery and

microphone. The data collected by the Narendra Modi app is being passed on to analytics company Clevertap. These data is further used for election campaign purpose and to provide the most contextual content to users [15].

11.3 Deep Fakes and Election Campaigns

Artificial intelligence refers to algorithms that mimic the intelligence of humans able to resolve the problem. Machine Learning refers to the algorithm, a subset of AI that includes a technique that enables the machine to improve at the task with experience. Advances in Artificial Intelligence (AI) and Machine Learning (ML) have enabled computer systems to create synthetic videos, a.k.a. deepfakes.

A deepfake video can show a person saying or doing something that they never said or did. An AI algorithm is taught, using deep learning, to manipulate actual video and audio to mislead viewers. Deepfake online videos are burgeoning as tools and applications to make fictional content cheaper, lowering barriers for non-experts to make them.

One of the ways to make deepfake videos involves encoding, decoding, and face-swapping. First, an AI encoder runs several face shots of two persons. Then, the encoder learns of similarities in the faces and compresses the images.

As the faces are different, one Artificial intelligence decoder picks the first person's face, and another decoder selects the second person's face. Then, the encoded images are fed into the "wrong" decoder to perform the face swap. This process is done on every frame to make a convincing video. Figure 11.4 shows the image of deepfake and impersonating examples of Barak Obama [16].

In some instances, Artificial intelligence is trained to create new images and videos from scratch using the Generative Adversarial Network (GAN). With the help of Machine Learning, GAN makes two neural networks to contest against each other in a game with a given training set. It helps generate new data and output related to the training set. In Figure 11.5 it has shown how AI brings Mona Lisa's looks from a different angle [17].

Artificial intelligence and Machine learning are now used in the election campaign to engage voters and be more informed about the political issue.

11.3.1 Deep Fake in Delhi Elections

Since its inception in late 2017, deepfake images and videos have primarily been used in pornography. Approximately 96% of all deepfakes online



Figure 11.4 An image of deep fake and impersonating examples of Barak Obama. Source: [16].

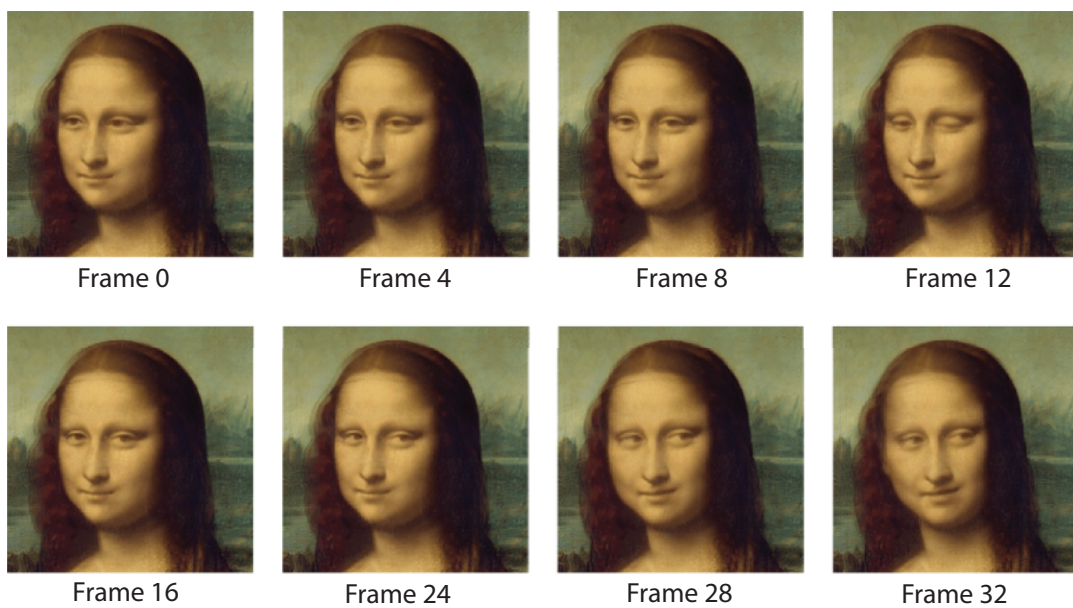


Figure 11.5 AI bring Mona Lisa's looks from different angle. Source: [17].

are non-consensual pornographic videos, according to Deep Trace Labs, a deepfake detection technology firm.

Deepfakes are now a growing concern across the world in recent years. Some of the most prominent targets have been celebrities, whose faces were mapped onto porn actors. Others have involved notable people such as Barack Obama name-calling US President Donald Trump and Facebook boss Mark Zuckerberg boasting about illegitimate data accessing—neither did that in real life.

An allegedly deepfake video of Gabon's President *Ali Bongo* led to a coup d'état, which was later crushed in the political sphere. Apart from this, Future Advocacy's deepfake video showed UK's Prime Minister *Boris Johnson* and Labour leader *Jeremy Corbyn* endorsing each other. The think-tank's uploaded a video in an attempt to show the potential of deepfakes in undermining democracy [18].

A day before the *Delhi* elections, two videos of *Bharatiya Janata Party (BJP)* Delhi unit President *Manoj Tiwari* goes viral. In this video, *Manoj Tiwari* urges citizens to vote for BJP in English and Haryanvi to 15 million voters via 5,800 WhatsApp groups. The videos were deepfakes, digital media firm *Vice* reported.

In the English video, *Mr. Tiwari* criticized his political opponent *Arvind Kejriwal* and asking voters to “please press the lotus button on 8 February to form *Modi*-led government.” However, in the original video in the Hindi language, *Mr. Tiwari* spoke about a completely different issue. The original video was mocked and audio overlaid in English and Haryanvi language. The lip-synching was tweaked to make it seem as if *Tiwari* was indeed speaking [19].

11.4 Social Media Bots

“A social bot is an automatic program that simulates human behavior on social networks. Social bots take part in discussions on Twitter or Facebook and act like human users. They spread content on a particular topic on social media—mostly to influence people's opinions.”

Bots now have a believable online profile with more advanced conversational skills and appear genuine users embedded in the human network. Humans partially manage some automated accounts. These profiles are known as sock puppets.

Social media bots are now everywhere on social media like commercial activity, counterterrorism, terrorism, Entertainment, Harassment, hate speech, information operation, notification, Social and civil engagement. Illustrations are depicted in Figure 11.6.

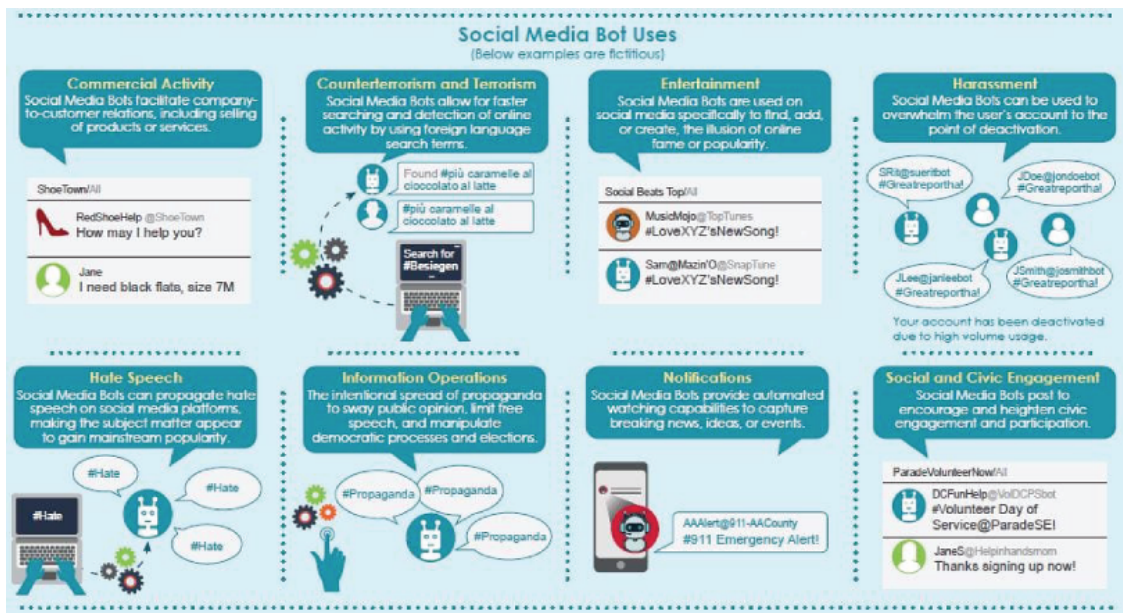


Figure 11.6 Social media bot uses. Source: [20].

Social media bots behaviors can identify commonly at three levels.

- Congregate of Bots*—Social media bots often congregate together and act randomly, making them easier to identify.
- Content Level*—Social media bots tend to use emoticons, exclamation points, or other content in more regular patterns than human uses on social media.
- Activity level*—Social media bots often have a higher level of activity (typically automated social media bots) compared to human social media behavior.

Details are presented in Figure 11.7.

Bots strategically amplify particular messages by spreading propaganda and shapes popular narratives. It attempts to set online political agendas, manipulate the public, and influence political processes. Bots, post, like, and dislike shares 24 × 7 h. By liking and retweeting, bots make topics and

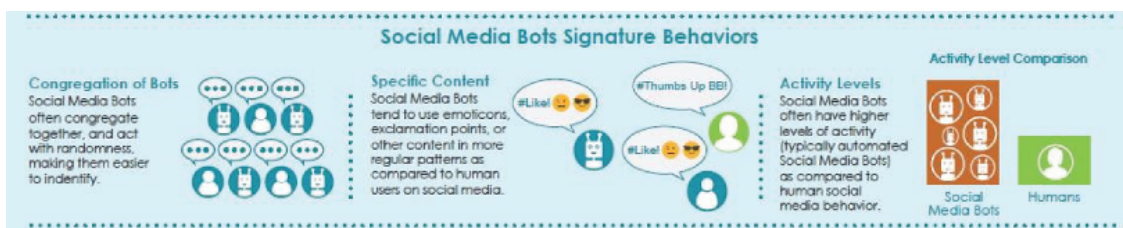


Figure 11.7 Social Media Bots Signature Behaviors. Source: [20].

opinions appear essential. Bots functioned as it is programmed. Bots are highly used for Political purposes to influence public discourse, military conflicts, and set agendas in elections [21].

A study claimed, “just six percent of bots on Twitter accounted for 31 percent of bad information”. Social media bots take a few seconds for bots to spread misinformation. Visualization of bots spread misinformation is depicted in Figure 11.8.

During the 2016 US Presidential election, nearly 400 thousand Twitter accounts, approximately 15% of the US’s entire Twitter population, were bots. These accounts were responsible for creating traffic in around 20% of online conversations.

In the UK, the research found bots play a strategic role in shaping discourse on Twitter about the polarizing BREXIT referendum. Hastag# associated with leaving the EU dominated the conversation. Less than 1% of accounts are responsible for generated one-third of all messages.

In the US, during Covid19 lockdown. Nearly half of the Twitter accounts discussed reopening the US have come from bots. Other reports highlighted in the top 50 influencer retweeter, 82% were bots, and in the top 1,000 retweets, 62% were bots [23].

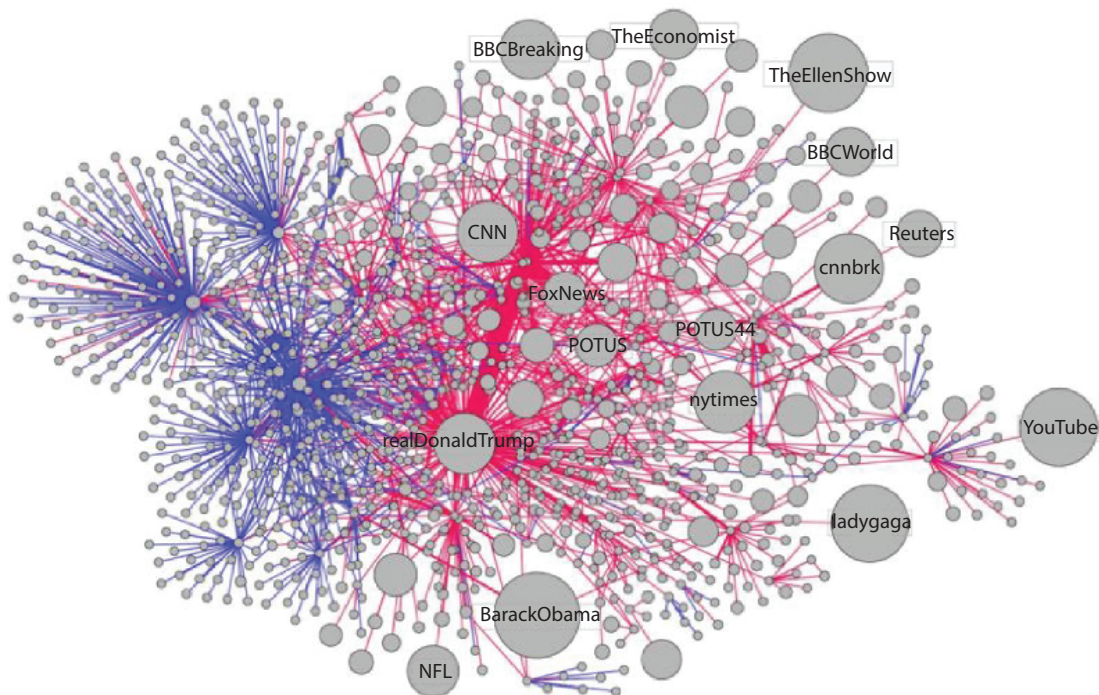


Figure 11.8 Visualization of the spread through social media of an article falsely claiming 3 million illegal immigrants voted in the 2016 presidential election. Source: [22].

11.5 Future of Artificial Intelligence and Machine Learning in Election Campaigns

AI and ML are the products of the advancement in technology. The future of AI and ML depends on technology and the internet around the world. Technology is getting advance every day. Internet is getting upgraded in higher generation speed. The effects of advancement in technology and the internet have been noticed in some recent election campaigns in developed and developing countries.

Now technology has become very easy to learn and required no expertise in operation. Thousands of software applications are freely available online. Online software is enabling to edit audio, video, and image. Unique features allow users to merge, mock, and infuse special effects in the video. These videos and audio can be circulated to millions in seconds across the globe with an internet connection. Now online campaigns are shaping citizen electoral choice. Soon, low-cost internet and technology advancement will shift the election campaign entirely online around democratic countries [24].

The application of AI and ML in election campaigns across democratic countries will increase in the future. A customized software application, Deepfake video, and bots will replace the door-to-door campaign to a large extent. A negative campaign will dominate, and other countries' involvement in the election will increase. The customized election campaign will use text messages and videos to target voters at an individual level. There is a high probability that voters' details will no more remain personal soon. Election management companies will access all the details of voters through a customized software application. However, the election will only be limited to some boundaries, but advertisement and election campaigns are now boundaryless and can be managed anywhere inside and outside the country. In a nutshell, the future of the election campaign will become data-driven. Citizens will elect their representatives for the house, who have used more advanced AI and ML in their election campaign.

References

1. Jain, B., Lok Sabha 2019: More than 90 crore voters register to vote. Times of India, New Delhi, 2019. Retrieved 10 November 2020, from http://timesofindia.indiatimes.com/articleshow/68620296.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.

2. Parveen, N. and Safiullah, M., Twitter and radio indicators of election outcomes: A study of Indian elections. *Int. J. Econ. Bus. Res.*, 22, 2–3, 278–289, 2021.
3. Safiullah, M., Pathak, P., Singh, S., The impacts of social media and news media on political marketing: An empirical study of 2014 Indian General Election. *Int. J. Bus. Excell.* (In-Press), 2021.
4. Safiullah, M., Primetime news coverage and electoral harvest—A study of 2014 Indian general election. *Int. J. Business Forecasting Market. Intell.*, 5, 4, 424–432, 2019.
5. Rodriguez, F., Scheinker, D., Harrington, R.A., Promise and perils of big data and artificial intelligence in clinical medicine and biomedical research. *Circ. Res.*, 123, 12, 1282–1284, 2018.
6. Safiullah, M., Pathak, P., Singh, S., Emergence of Social Media and its implications for Public Policy: A study of Delhi Assembly Election 2013. *Manage. Insight*, 12, 1–6, 2016.
7. Safiullah, M., Pathak, P., Singh, S., Anshul, A., Social media as an upcoming tool for political marketing effectiveness. *Asia Pac. Manage. Rev.*, 22, 1, 10–15, 2017.
8. Smith, J., Big data analytics—Decisions, TDWI, Europe, 2020, Retrieved 30 November 2020, from <https://www.autocab.com/wp-content/uploads/2016/09/big-data-analytics-decisions-1140x682.jpg>.
9. Jha, Want to Win an Election? Use AI And Machine Learning, Greatlearning, India, 2020, Retrieved 10 November 2020, from <https://www.mygreatlearning.com/blog/how-ai-and-machine-learning-can-win-elections/>.
10. Cadwalladr, *et al.*, How Trump Consultants Exploited the Facebook Data of Millions, The New York Times, London, 2018, Retrieved 10 November 2020, from <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html>.
11. Krishanmurthy, R., The Role of Data in Elections, Analyticsteps, India, 2020, Retrieved 10 November 2020, from <https://www.analyticssteps.com/blogs/role-data-elections>.
12. Rivera, Donald Trump's 2020 campaign app gets a boost with Phunware partnership, one day after Twitter fact-check, USA TODAY, USA, 2020, Retrieved 10 November 2020, from <https://www.usatoday.com/story/tech/2020/05/27/donald-trump-reelection-app-phunware/5271268002/>.
13. Markowitz, E., AI, data gathering to play pivotal role in 2020 elections, IOTTimes, USA, 2020, Retrieved 10 November 2020, from <https://iot.eetimes.com/ai-data-gathering-to-play-pivotal-role-in-2020-elections/>.
14. CNBC, *How Presidential Campaigns Use Apps for Data Collection* [Video], CNBC, USA, 2020, Retrieved from <https://youtu.be/-tzyDB48LDQ?t=456>.
15. Mishra, P.K., The Narendra Modi app: The secret weapon in BJP's elections arsenal, The internet centre for & society, India, 2018, Retrieved 30 November 2020, from <https://archive.factorodaily.com/narendra-modi-app-bjp-2019-election/>.

16. Knight, W., A new deepfake detection tool should keep world leaders safe—For now, MIT Technology Review, USA, 2020, Retrieved 30 November 2020, from <https://www.technologyreview.com/2019/06/21/134815/a-new-deepfake-detection-tool-should-keep-world-leaders-safe-for-now/>.
17. Goswami, A., Animating The Mona Lisa Effect With Deep Learning Using Tensorflow.js, MarkTechPost, 2020, Retrieved 30 November 2020, from <https://www.marktechpost.com/2020/10/03/animating-the-mona-lisa-effect-with-deep-learning-using-tensorflow-js/>.
18. Bienkov, Boris Johnson appears to endorse Jeremy Corbyn for prime minister in a convincing deepfake video, Business Insider, India, 2019, Retrieved 10 November 2020, from <https://www.businessinsider.in/politics/news/boris-johnson-appears-to-endorse-jeremy-corbyn-for-prime-minister-in-a-convincing-deepfake-video/articleshow/72022620.cms>.
19. Xavier, J., Deepfakes enter Indian election campaigns, The Hindu, India, 2020, Retrieved 10 November 2020, from <https://www.thehindu.com/news/national/deepfakes-enter-indian-election-campaigns/article30880638.ece>.
20. Ruiz, A., New Cybersecurity Series: PROTECT IT.—Social Media Bots—THE SAFE, The Safe, US, 2020, Retrieved 30 November 2020, from <https://sites.northwestern.edu/thesafe/2019/10/21/new-cybersecurity-series-protect-it-social-media-bots/>.
21. Marwick, A. and Lewis, R., *Media manipulation and disinformation online*, pp. 7–19, Data & Society Research Institute, New York, 2017.
22. Ouellette, J., Study: It only takes a few seconds for bots to spread misinformation, Ars Technica, USA, 2020, Retrieved 30 November 2020, from <https://arstechnica.com/science/2018/11/study-it-only-takes-a-few-seconds-for-bots-to-spread-misinformation/>.
23. Swain, How robots are coming for your vote, BBC, UK, 2019, Retrieved 10 November 2020, from <https://www.bbc.com/future/article/20191108-how-robots-are-coming-for-your-vote>.
24. Safiullah, M., Pathak, P., Singh, S., Anshul, A., Social media in managing political advertising: A study of India. *Pol. J. Manage. Stud.*, 13, 2, 121–130, 2016.