1. Develop a problem statement

In Kenya, social media is fundamentally reshaping how citizens communicate and how brands and campaigners get their message across news media and social media the information is spread high speed but without accuracy and hence detection mechanism should be able to predict news fast enough to tackle the dissemination of fake news, Such misinformation can generate elusive thoughts and opinions, collective hysteria or other serious consequences. Therefore, automated disinformation detection is now needed in this area of research. Text classification problems related to detecting deceptive, misleading, or fraudulent information have been applied to use cases such as spam filtering, phishing alerts, online reviews and opinion spam, and fake social profiles; in this research, we are going to combine machine learning with knowledge engineering to detect disinformation in social network.

2. Identify data sources to help answer the problem

Facebook, News, Newspapers, Twitter, Websites and Communication Authority of Kenya datasets

3. Indicate the tools and approach you would use

In determining whether a news article is fake or legitimate is an example of a binary classification problem. Classification is one of two supervised machine-learning problems, the other being regression. Classifiers take data as input and assign a label as output

Therefore, to address this problem we will employ Machine Learning (Naïve Bayes, Support Vector Machine,), Deep Learning (PyTorch),Natural Language Processing algorithms (NLTK) and Performance evaluation (Confusion matrices).

Approach:

Machine Learning Workflow:

The disinformation classification problem can be summarized as follows:

Gather annotated dataset=> à Perform necessary pre-processing (stemming/lemmatization, stop word removal, normalization) => Feature selection and extraction à Split data into training and test sets (with optional validation sets) à=>Train classifier using training set => Test performance using test data set.

Web scraping using machine learning libraries (Beautiful Soup, Scrapy and Selenium)

Analyzing live streaming tweets using machine learning libraries (Tweepy, Spacy)