You've won a 2-million-dollar worth lottery' we all get such spam messages. How can AI be used to detect and filter out such spam messages?

SOL

To understand spam detection, let's take the example of Gmail. Gmail makes use of machine learning to filter out such spam messages from our inbox. These spam filters are used to classify emails into two classes, namely spam and non-spam emails.

Let's understand how spam detection is done using machine learning:
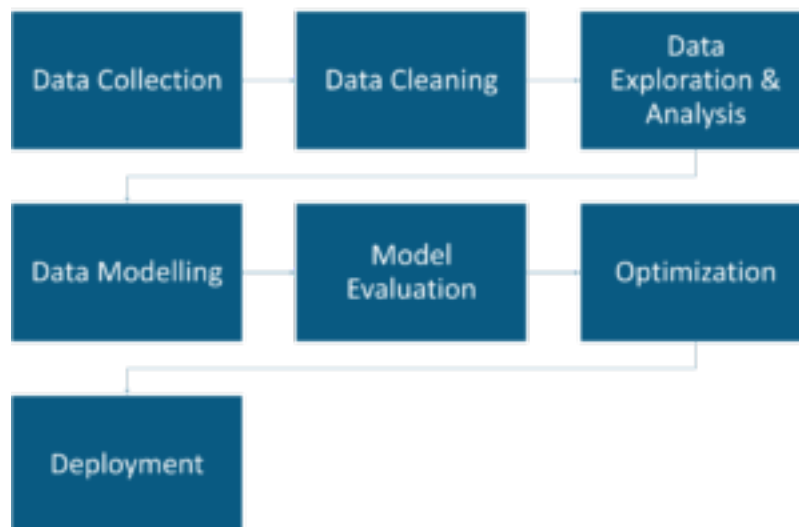
- A machine learning process always begins with data collection. We all know the data Google has, is not obviously in paper files. They have data centers which maintain the customer's data. Data such as email content, header, sender, etc are stored.
- This is followed by data cleaning. It is essential to get rid of unnecessary stop words and punctuations so that only the relevant data is used for creating a precise machine learning model. Therefore, in this stage stop words such as 'the', 'and', 'a' are removed. The text is formatted in such a way that it can be analyzed.

- After data cleaning comes data exploration and analysis. Many a time, certain words or phrases are frequently used in spam emails. Words like "lottery", "earn", "full-refund" indicate that the email is more likely to be a spam one. Such words and co-relations must be understood in this stage.
- After retrieving useful insights from data, a machine learning model is built. For classifying emails as either spam or non-spam you can use machine learning algorithms like Logistic Regression, Naïve Bayes, etc. The machine learning model is built using the training dataset. This data is used to train the model and make it learn by using past user email data.
- This stage is followed by model evaluation. In this phase, the model is tested using the testing data set, which is nothing but a new set of emails. After which the machine learning model is graded based on the accuracy with which it was able to classify the emails correctly.
- Once the evaluation is over, any further improvement in the model can be achieved by tuning a few variables/parameters. This stage is also known as parameter tuning. Here, you basically try to improve the efficiency of the machine learning model by tweaking a few parameters that you used to build the model.
- The last stage is deployment. Here the model is deployed to the end users, where it processes emails in real time and predicts whether the email is spam or non-spam.