

PROJECT PROPOSAL

TEXT MESSAGE SPAM CLASSIFIER APP

CSE-5334 – 001

Name: Kapish Shewale

UTA ID: 1001870794

These days, with increase in cyber crimes and phishing attacks, it is very important that one can identify spam text messages. Everyone these days carries a cell phone and are connected to their social circle via text messages. This makes the text messages a good attack surface.

This is where the 'Text Message Spam Classifier App' comes to the rescue. A user can simply input the text message to this app and this app will automatically tell weather the message is a spam or not spam.

The app is backed by Machine learning **Naïve Bayes Classifier**.

The user just has to input his text message into the text box and when checked, it will tell weather the message is a spam or not. For security reasons, the app does not store any text messages in its data bases.

This is currently a web-app, but as the **future development phases**, this app can be made into a mobile phone app and can automatically classify the text messages if given appropriate permissions. Only with the user permission **in the future**, this app can simultaneously train itself with user's received messages. Hence it will get better and better with passing time.

Some Popular apps using this feature already are:

1. Android text message app – “ Messages ”
2. Appple text message app – “ iMessage ”
3. True Caller
4. Gmail (**To classify mails**)
5. Outlook (**To classify mails**)

The dataset I will use is the **SMS Spam Collection Dataset** provided by **Kaggle** and originally by the **UCI Machine Learning Repository**.

The link for dataset is: <https://www.kaggle.com/uciml/sms-spam-collection-dataset>