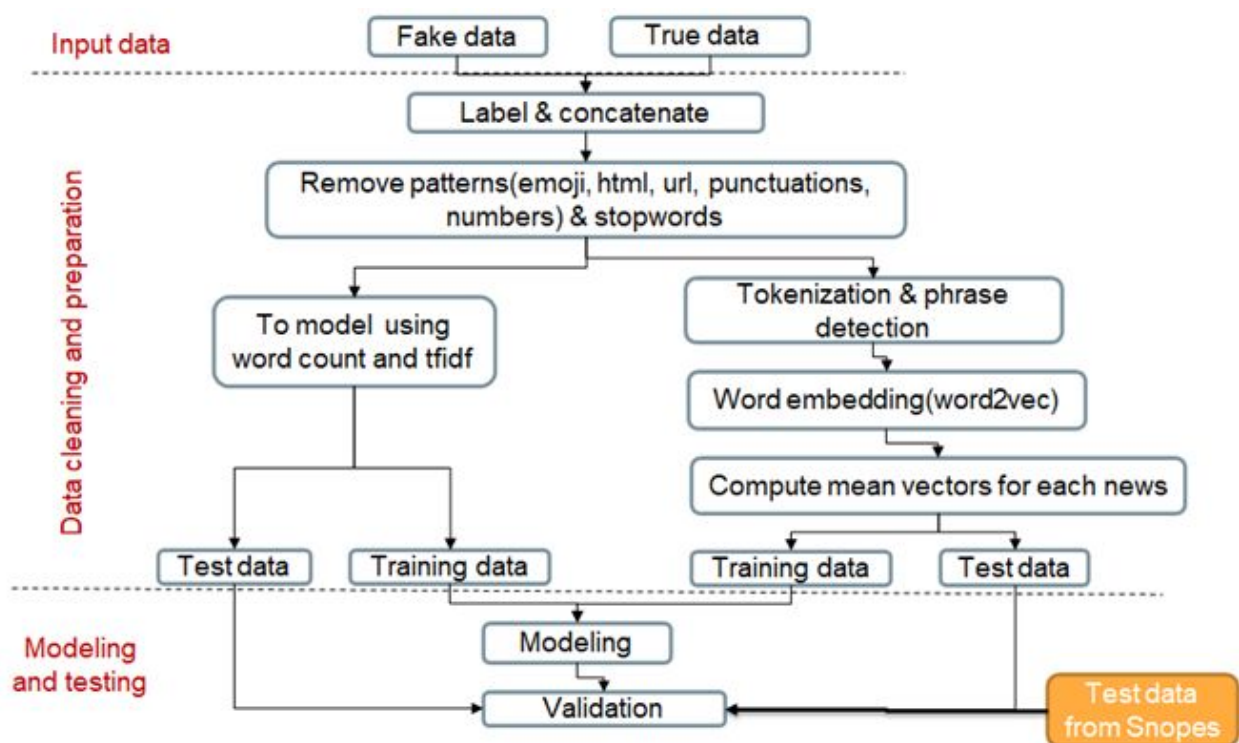


## Summary of the project

The objective of this project is to identify fake and real news based on its content using machine learning techniques. The result can help news agencies to identify fake news, to falsify it, broadcast the truth and play a role for society. This saves the community from confusion, wrong decision, misinterpretation and wrong judgment coming as a result of fake news. From the news agency's perspective, unless they fight fake news, at some point they might lose their audience.

The data used for this project is available on Kaggle (<https://www.kaggle.com/clmentbisailon/fake-and-real-news-dataset>) in two (.csv) files ; for fake and real news separately. Each dataset contains the news title, text, subject and date at which the article was posted. The fake news dataset contains 23,481 articles(rows) and the real news contains 21,417 articles(rows).

The approach to perform the analysis and modeling is summarized in the figure below.



### Important files in this folder

Description	File name
Notebook containing data exploration , wrangling and statistical test	<a href="#"><u>Data wrangling, cleaning and inferencial statistics.ipynb</u></a>
Notebook for modeling and testing	<a href="#"><u>Capstone project_2 modeling and testing.ipynb</u></a>
Capstone project-2 milestone report-2	<a href="#"><u>Capstone project -2 milestone report-2.docx</u></a>
Presentation	