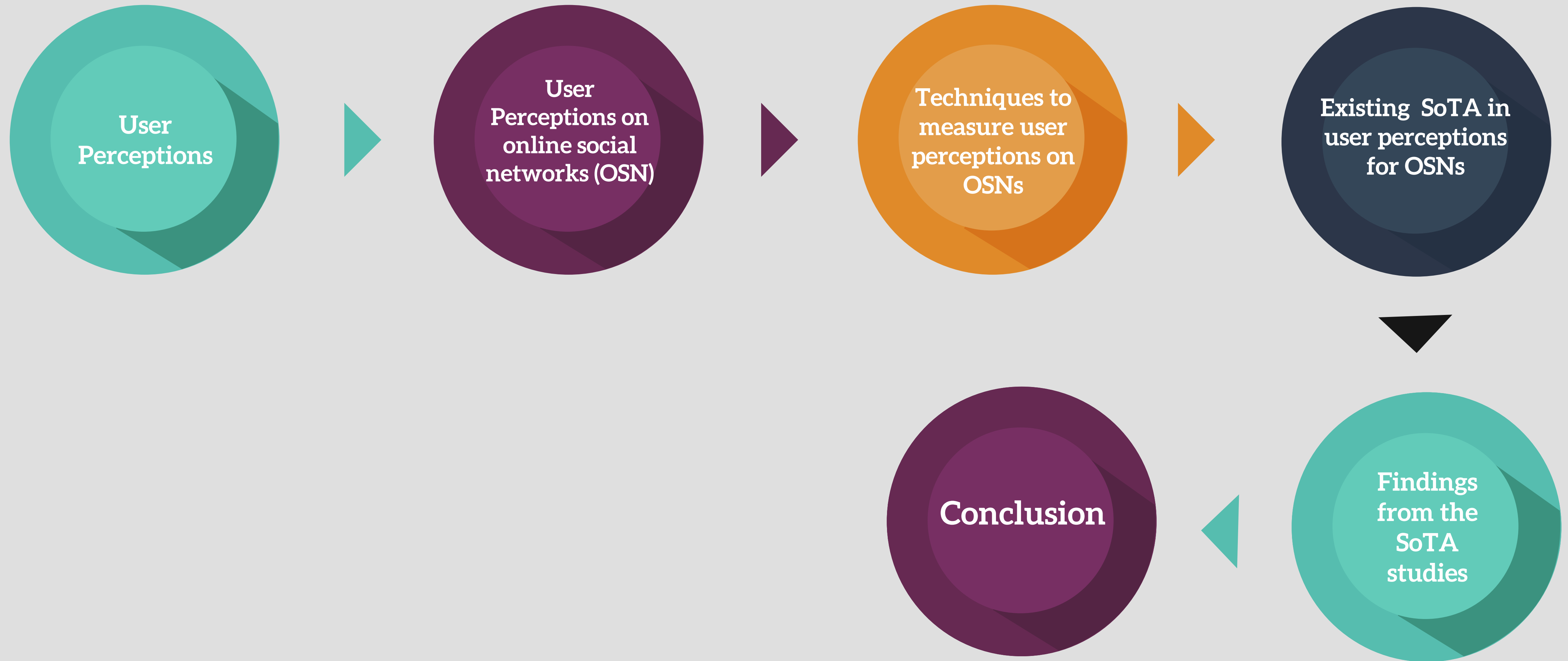


**computing**  
**conference 2022**

**A survey of artificial intelligence  
techniques for user perceptions'  
extraction from social media data**

**Presenter: Sarang Shaikh**

# Agenda of the talk




# User Perceptions

# User Perceptions

- Measuring and analyzing user perceptions and behaviors in order to make user-centric decisions has been a topic of research for a long time even before the invention of social media platforms.
- User perceptions about any events, topics, policies, etc. have always appealed the attention of policy and decision makers.
- These perceptions are always considered as strong evidence for making and adjusting user-centric decisions.
- The traditional method of analyzing/investigating user perceptions is usually based on data collection from survey polls and questionnaires.

# User Perceptions

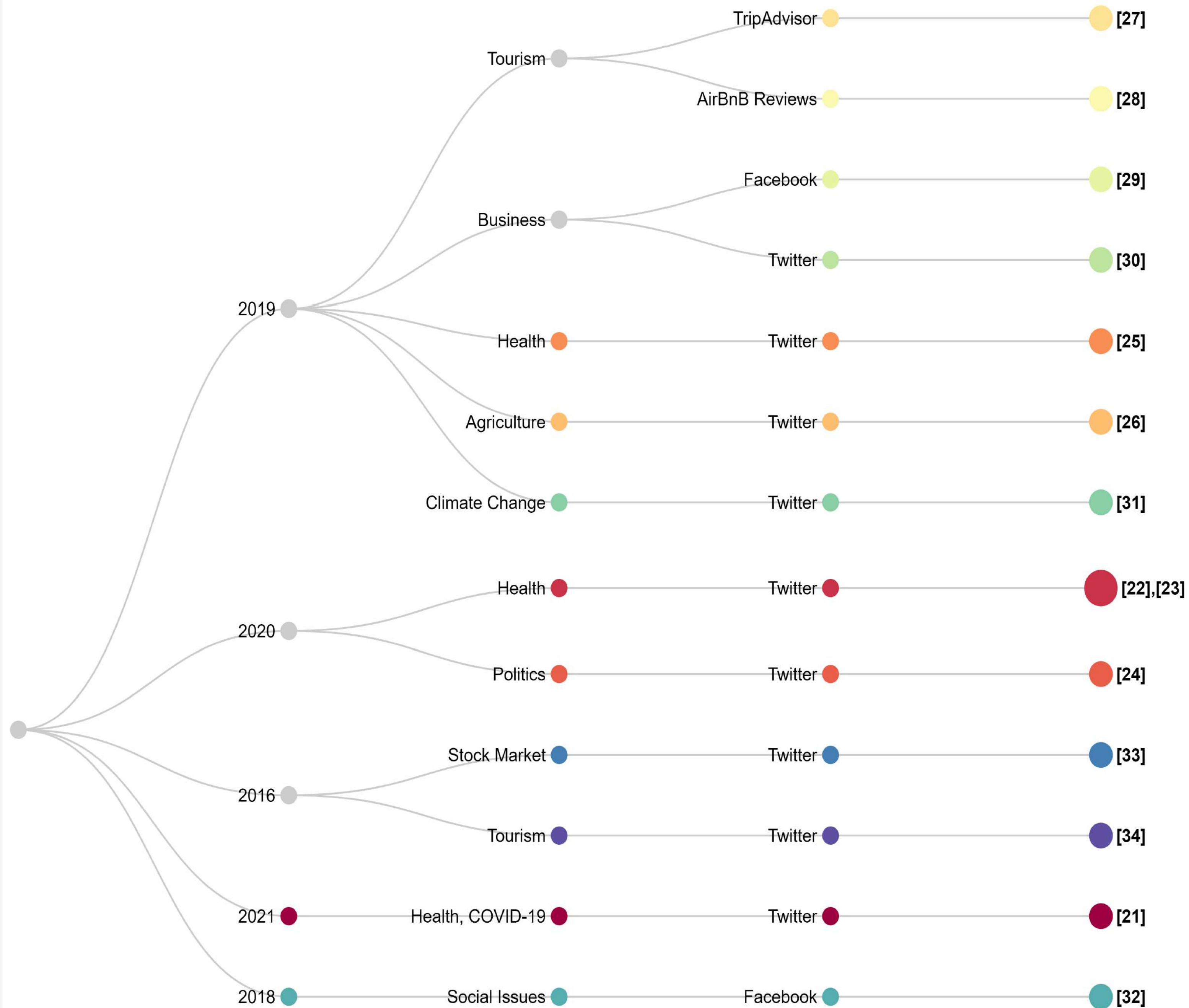
- However, some of the researchers have argued that these approaches more likely represent a small group of individual user perceptions rather than public user perceptions.
- Furthermore, due to the time and cost constraints involved in survey and questionnaire activities, the amount of collected data is very limited and hence it restricts the overall findings for understanding user perceptions to a large extent.



# User Perceptions on online social networks (OSN)

# User Perceptions on OSNs

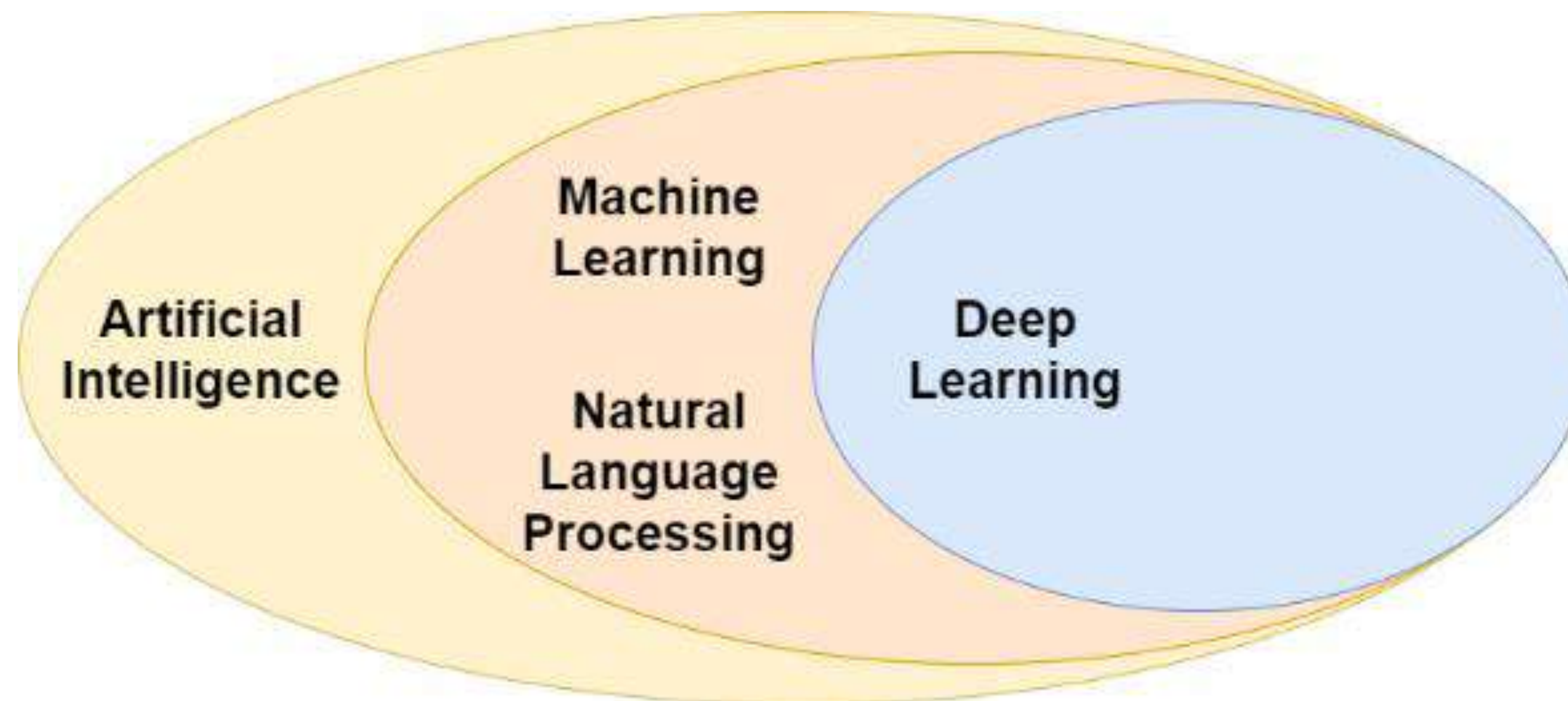
- Nowadays, social media platforms like Twitter, Facebook, LinkedIn, etc. offer a new way of understanding and measuring user perceptions.
- There has been an increase in adoption and use of social media platforms by the general public as well as the enterprises, industry owners, government officials, scientists, scholars, etc.
- The understanding and extraction of user perceptions from social media data has been widely studied in several domains including social science, education, politics, marketing, healthcare, finance and disaster management.



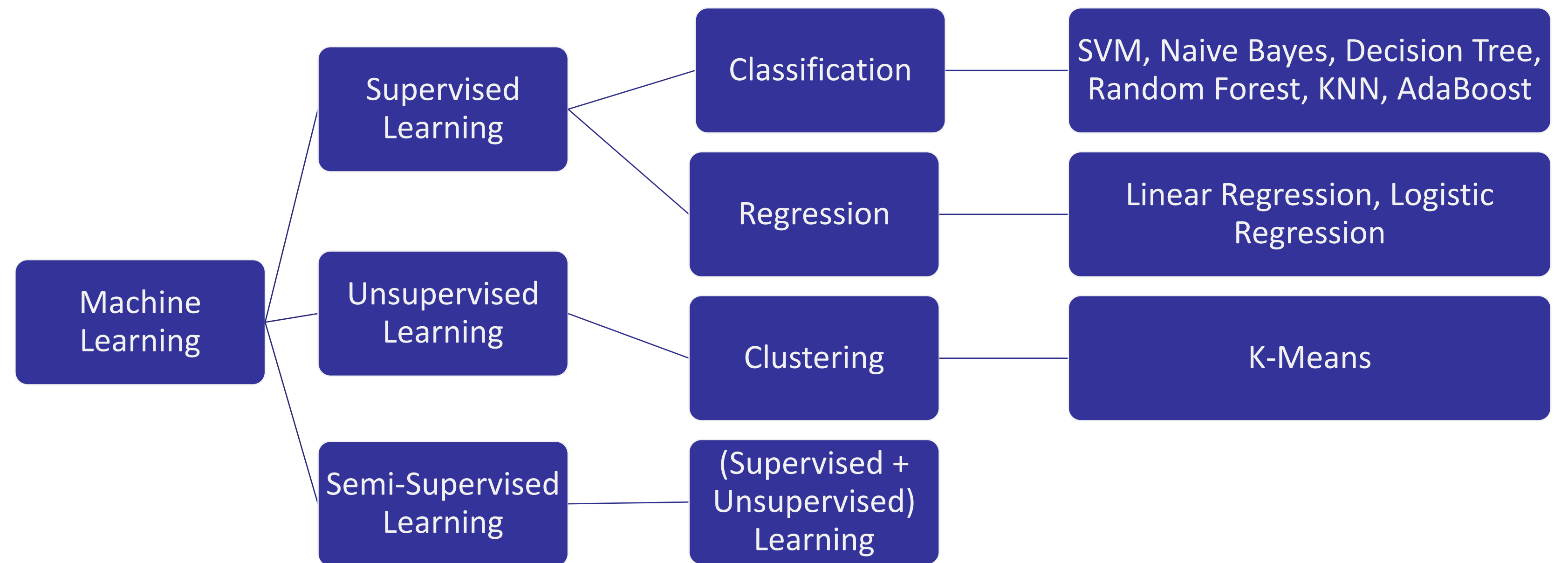
# Existing SoTA



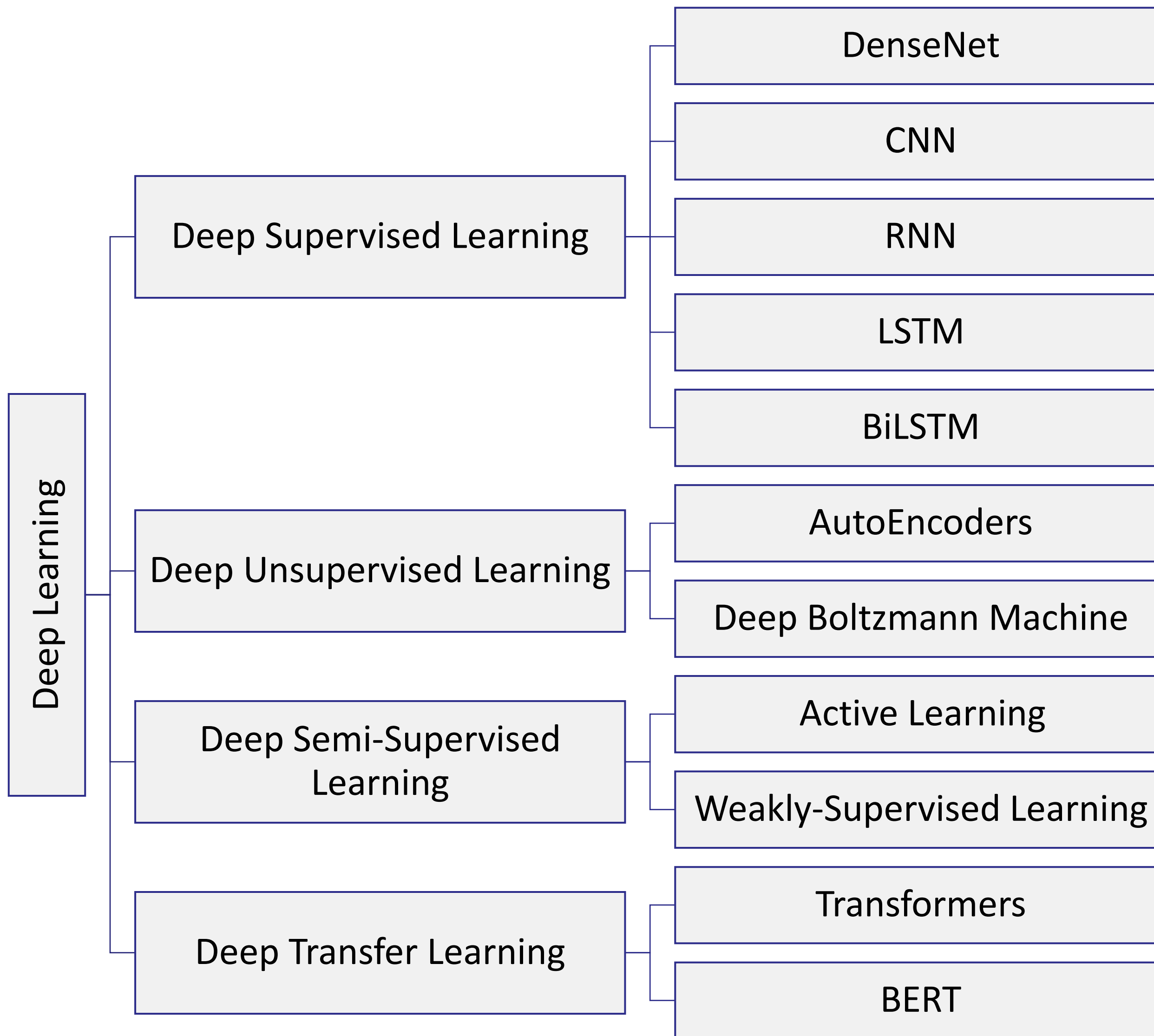
**Techniques to  
measure user  
perceptions  
analysis on  
OSNs**



# Taxonomy of Machine Learning techniques



# Taxonomy of Deep Learning techniques



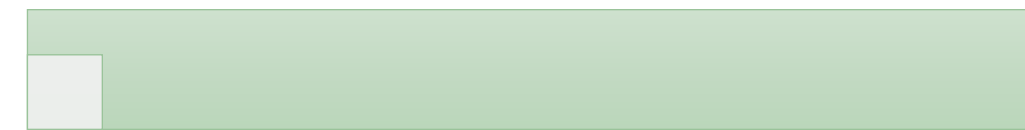
# Taxonomy of Natural Language Processing techniques

## Subjectivity Analysis



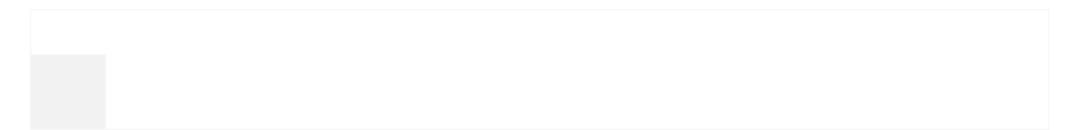
- Lexicon Based
- ML/DL Based

## Sentiment Analysis



- Lexicon Based
- ML/DL Based
- Transfer Learning Based

## Keyword/Keyphrase Extraction



- Topic Modelling Based
- Statistical Based
- ML/DL Based
- Graph Based
- Hybrid Approach

Existing SoTA studies for  
analysing user  
perceptions on online  
social networks

# Machine Learning techniques

#	Title	ML Model Used	Dataset Used	Social Media Platform Used	Application Area
1	Crowdsourcing and collaborative learning environments based on SM	Gaussian Naïve Bayes	Twitter, Facebook, LinkedIn	Twitter, Facebook, LinkedIn	Business Intelligence
2	Data analytic learning with strategic decision making	DT	Twitter hashtag, Meme tracker, and Yelp	Twitter	Business Intelligence
3	Fake profile detection	MRF	Facebook	Facebook	Crime detection
4	Cyberbullying Detection based on Semantic-Enhanced Marginalized Denoising Auto-Encoder	Bow, SVM, LDA	Twitter, Myspace	Twitter, Myspace	Crime detection
5	Identifying Epidemics	SVM, NB, and RF	Weibo		Epidemics
6	Detection of influenza epidemics	Linear Regression, Multiple Regression	Twitter	Twitter	Epidemics

# Deep Learning techniques

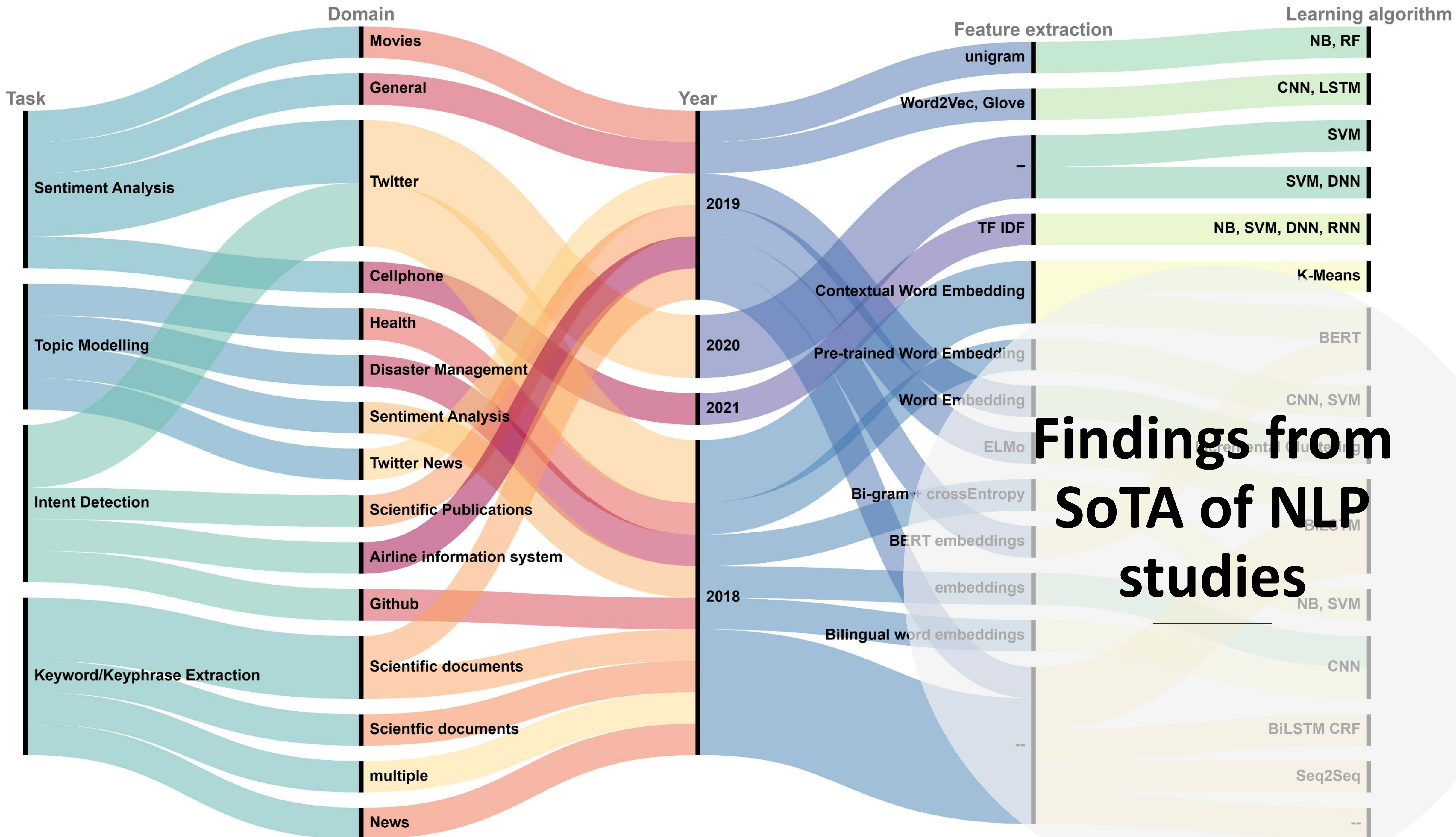
Title	DL Model Used	Dataset Used	Social Media Platform Used	Year	Application Area
<b>Deep Learning for Hate Speech Detection in Tweets</b>	FastText + CNN, LSTM	Racist tweets dataset	Twitter	2017	Hate Speech
<b>Detecting Offensive Language in Tweets Using Deep Learning</b>	RNN	Hate speech tweets	Twitter	2018	Hate Speech
<b>Multi-layers Convolutional Neural Network for Twitter Sentiment Ordinal Scale Classification</b>	CNN	SemEval challenge dataset	Twitter	2018	Sentiment Analysis
<b>Bloom's Learning Outcomes' Automatic Classification Using LSTM and Pretrained Word Embeddings</b>	FastText + LSTM	Course learning outcomes dataset	Twitter	2021	Bloom's Taxonomy
<b>Evaluating Polarity Trend Amidst the Coronavirus Crisis in Peoples' Attitudes toward the Vaccination Drive</b>	FastText + LSTM	Covid-19 tweets	Twitter	2021	Sentiment Analysis
<b>Deep learning-based personality recognition from text posts of online social networks</b>	CNN, RNN	Facebook posts	Facebook	2018	Personality Recognition
<b>Personality recognition from Facebook text for Portuguese language</b>	LSTM	Facebook posts	Facebook	2018	Personality Recognition

# Natural Language Processing techniques

Reference	Type	Feature extraction	Learning Algorithm	Domain	Year	Task
<b>Implementation of sentiment classification of movie reviews by supervised machine learning approaches</b>	Supervised	unigram	NB, RF	Movies	2019	Sentiment Analysis
<b>Evaluation of deep learning techniques in sentiment analysis from twitter data</b>	Supervised	Word2Vec, Glove	CNN, LSTM	General	2019	Sentiment Analysis
<b>Deep representation learning for clustering of health tweets</b>	Unsupervised	Contextual Word Embedding	K-Means	Health	2018	Topic Modelling
<b>Tweets classification with bert in the field of disaster management</b>	Supervised	Contextual Word Embedding	BERT	Disaster Management	2018	Topic Modelling
<b>Exploiting topic-based adversarial neural network for cross-domain keyphrase extraction</b>	Unsupervised	--	BiLSTM	Scientific documents	2018	Keyword/ Keyphrase Extraction
<b>Bidirectional lstm recurrent neural network for keyphrase extraction</b>	Unsupervised	--	BiLSTM	Scientific documents	2018	Keyword/ Keyphrase Extraction

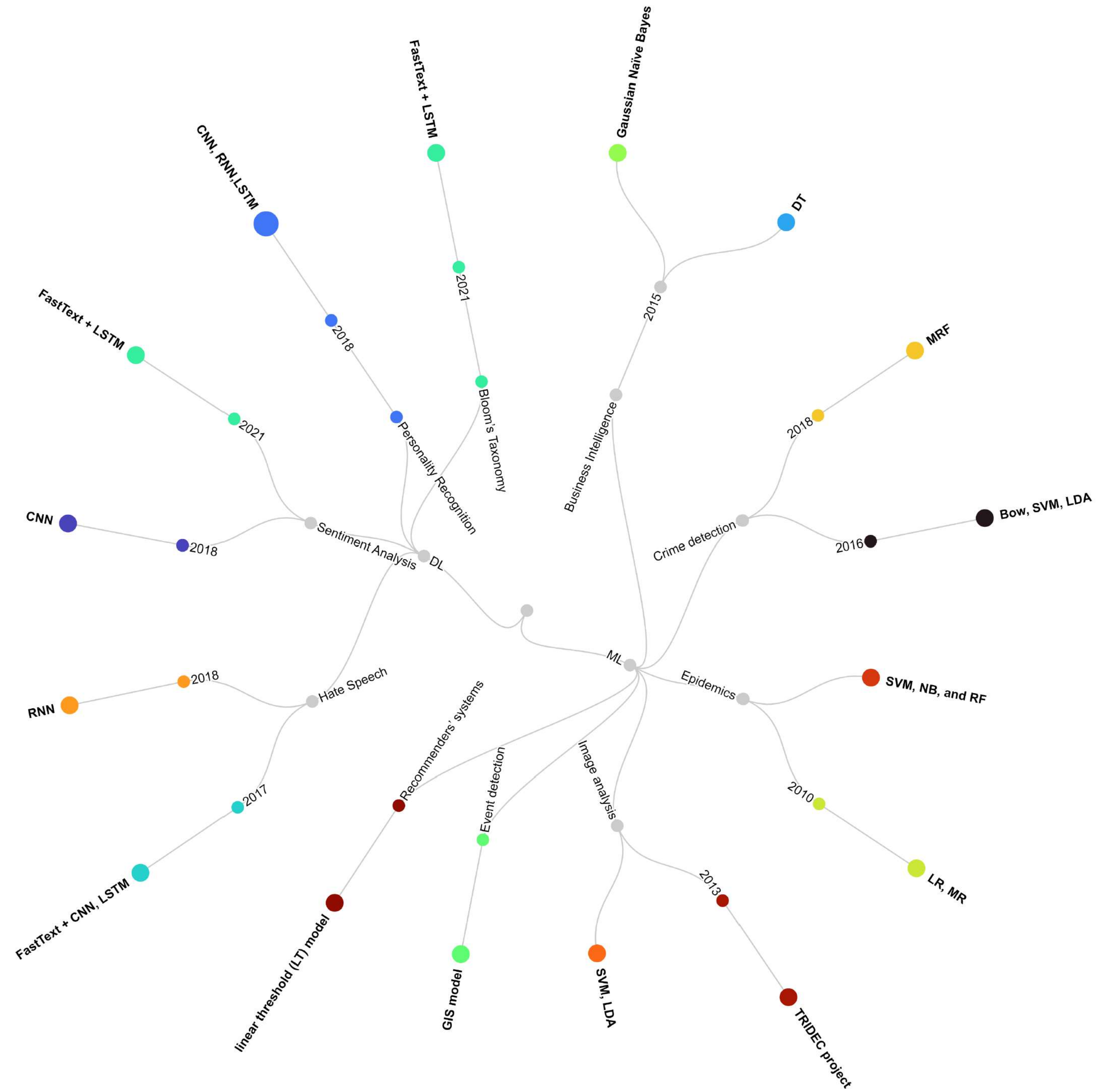


# Findings from the SoTA studies



# Findings from SoTA of NLP studies

# Findings from SoTA of ML/DL studies



# Conclusion

- In this paper, we described several artificial intelligence techniques including machine learning, deep learning and natural language processing in detail for the purpose of social media data analysis.
- Along with describing the techniques, we also conducted state-of-the-art review of studies where these techniques are applied for social media data analysis.
- The findings of review resulted in the identification of existing domains where the techniques are most widely used.
- The identified domains are health, disaster management, online social networks, news and scientific documents.
- Also, the review revealed the shift of the learning algorithms from ML to DL in recent studies in the year 2021.
- Overall, the review provides good discussions related to the different type of algorithms applied in various domains for achieving various types of tasks.

# Contact Us

Sarang Shaikh, Sule Yildirim Yayilgan, Erjon Zoto,  
Mohamed Abomhara

Dept. of Information Security and Communication  
Technology, Norwegian University of Science and  
Technology (NTNU), Gjøvik, Norway.

[sarang.shaikh@ntnu.no](mailto:sarang.shaikh@ntnu.no), [sule.yildirim@ntnu.no](mailto:sule.yildirim@ntnu.no),  
[erjon.zoto@ntnu.no](mailto:erjon.zoto@ntnu.no), [mohamed.abomhara@ntnu.no](mailto:mohamed.abomhara@ntnu.no)