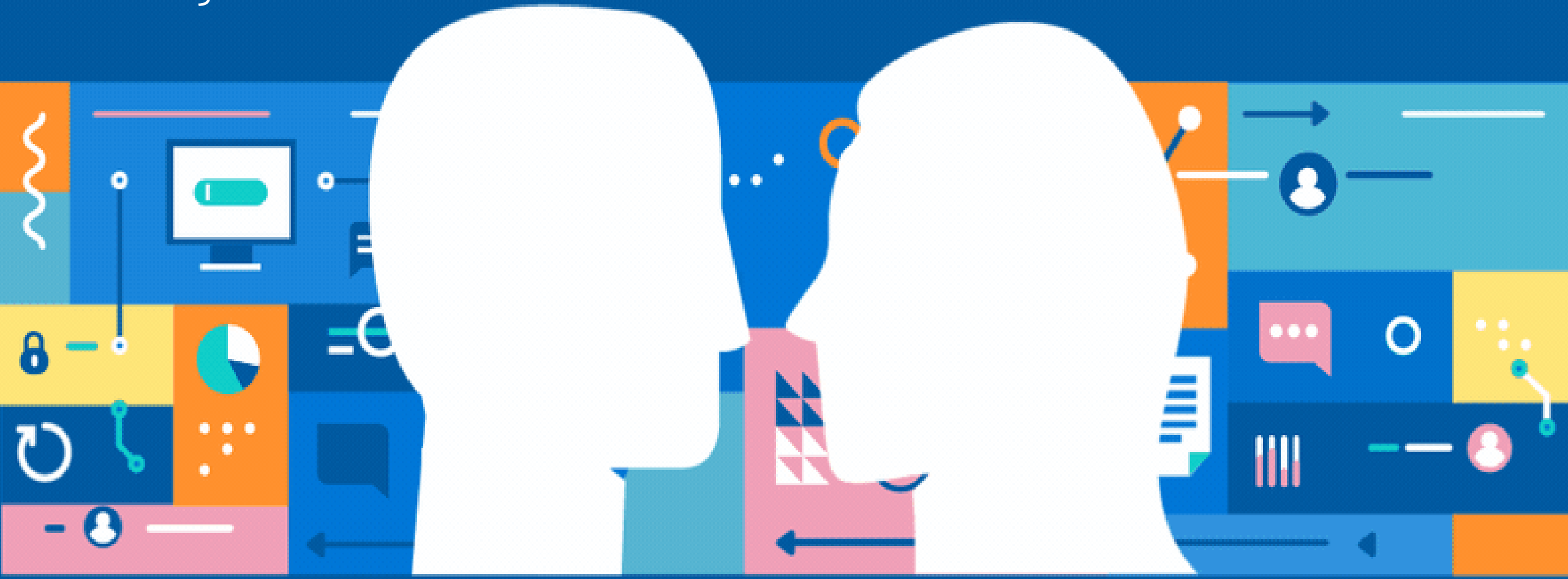
Sentiment Analysis Using Transfer Learning

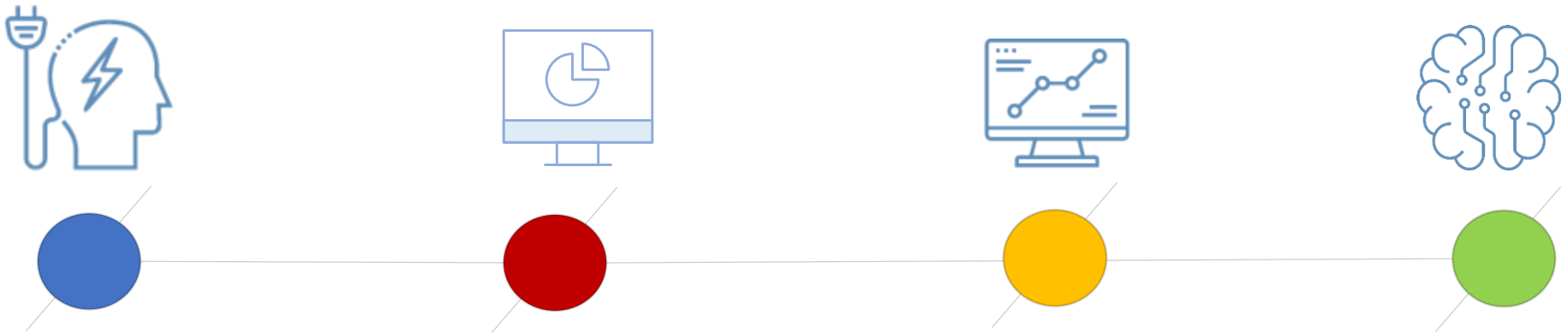
Santoshi Tadanki
Tamoghna Saha



Agenda

- **Evolution of business decisions**
- Need to process natural language
- Applications of NLP
- Process of natural language comprehension
- Pre-processing techniques for natural language
- Sentiment Analysis and Transfer Learning
- Annexure

Evolution of Business Intelligence



Executive information systems

On premise storage

Real time analysis

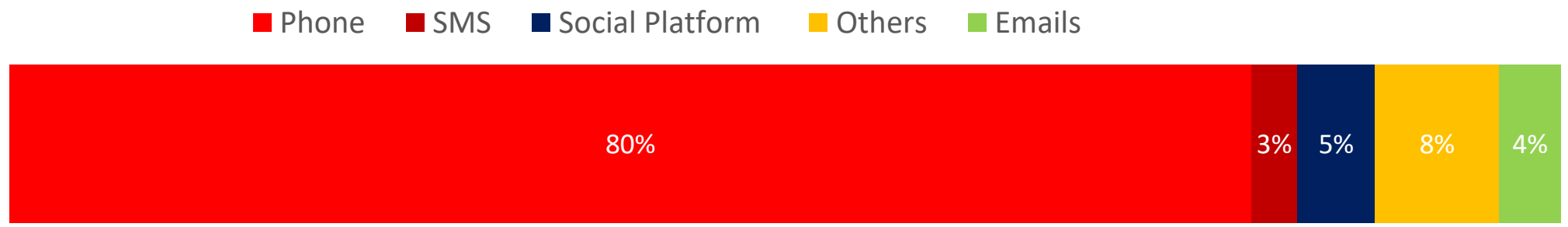
Intelligence

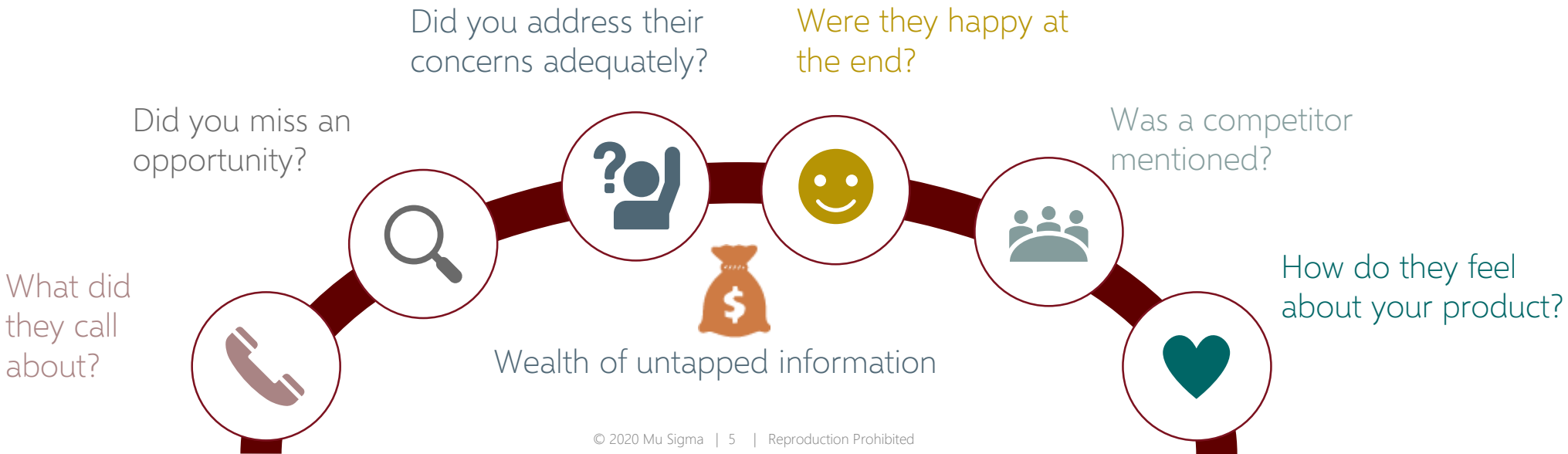
Agenda

- Evolution of business decisions
- **Need to process natural language**
- Applications of NLP
- Process of natural language comprehension
- Pre-processing techniques for natural language
- Sentiment Analysis and Transfer Learning
- Annexure

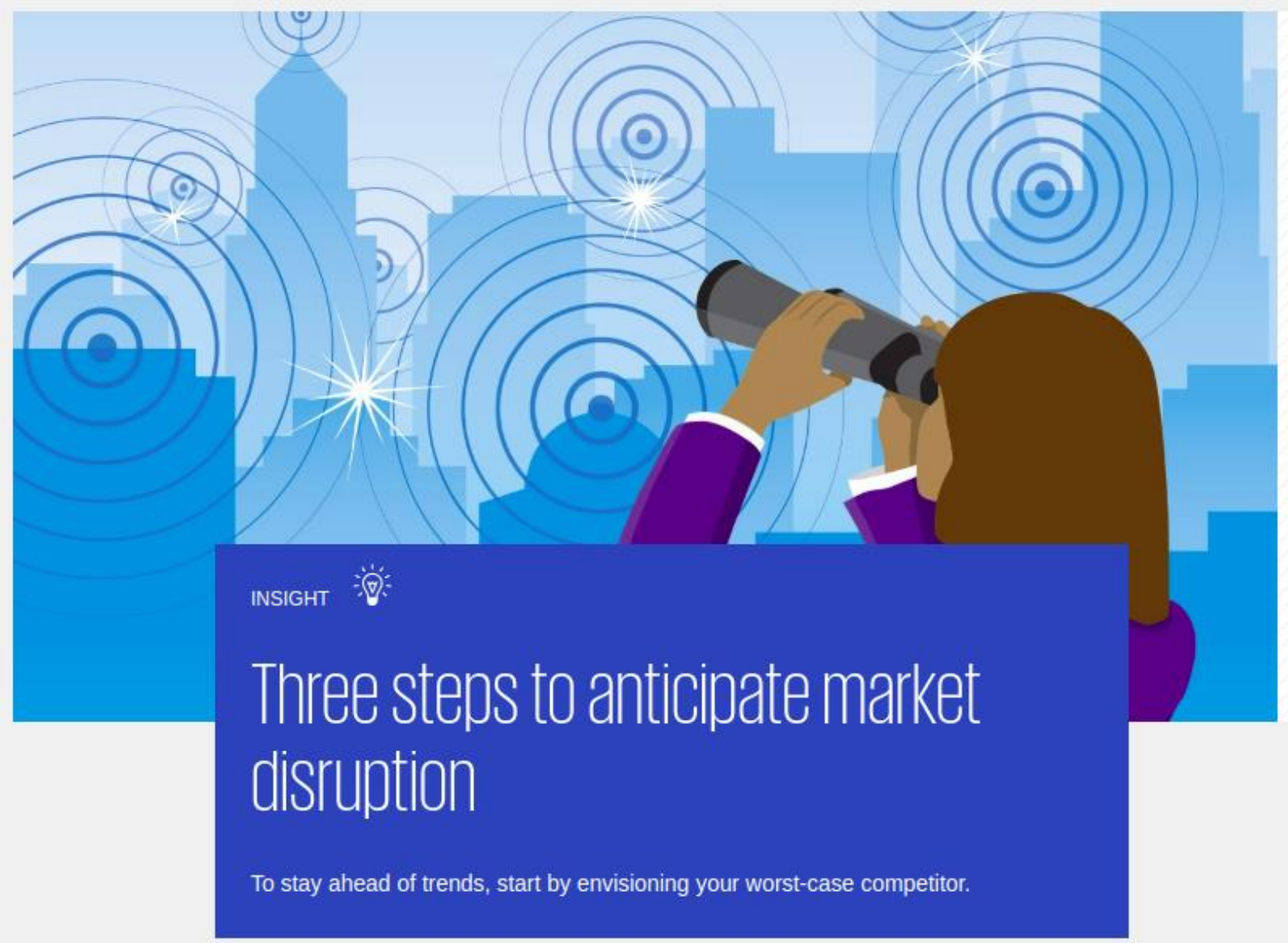
The majority of the customer touch points after purchase are unstructured

Final customer touch point





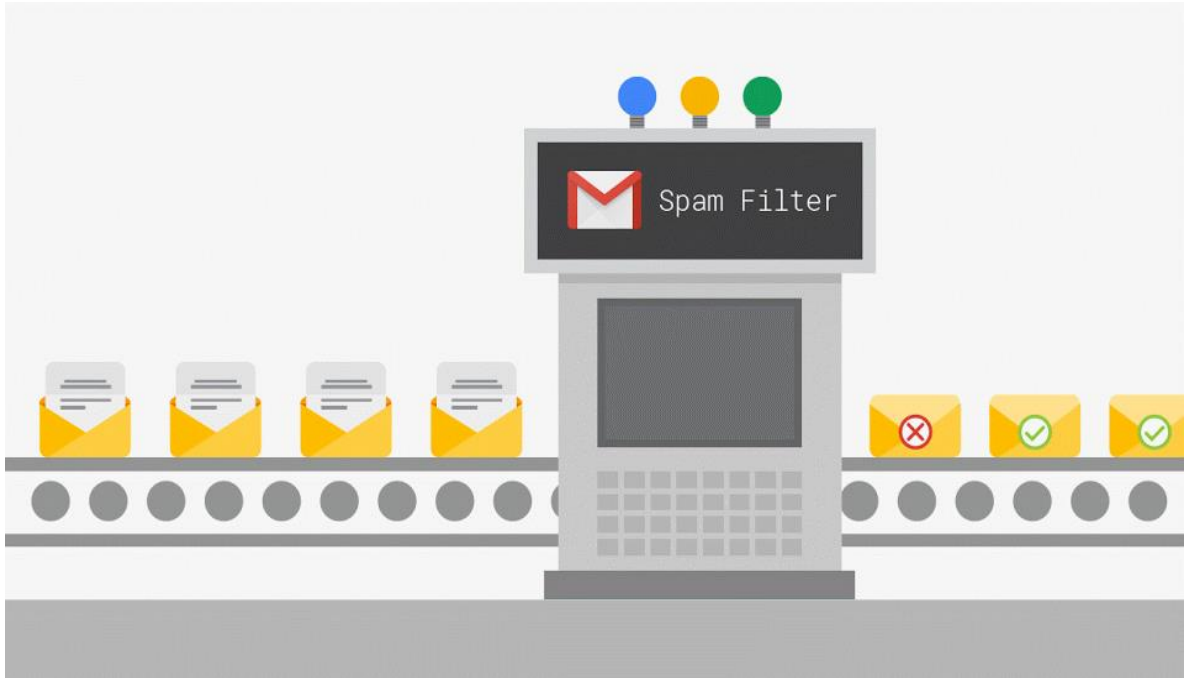
One of the big 4 say that to identify market disrupters, we should take an outside-in approach

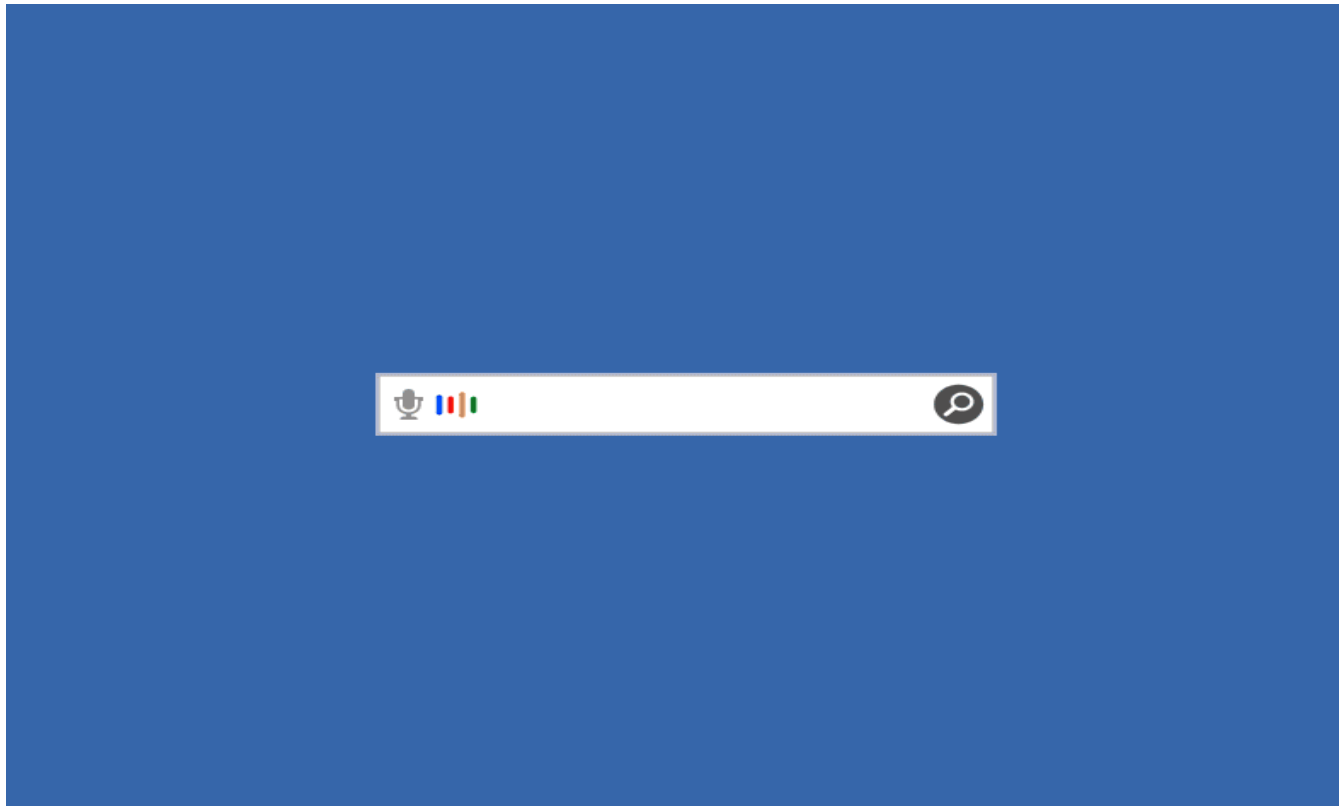


Agenda

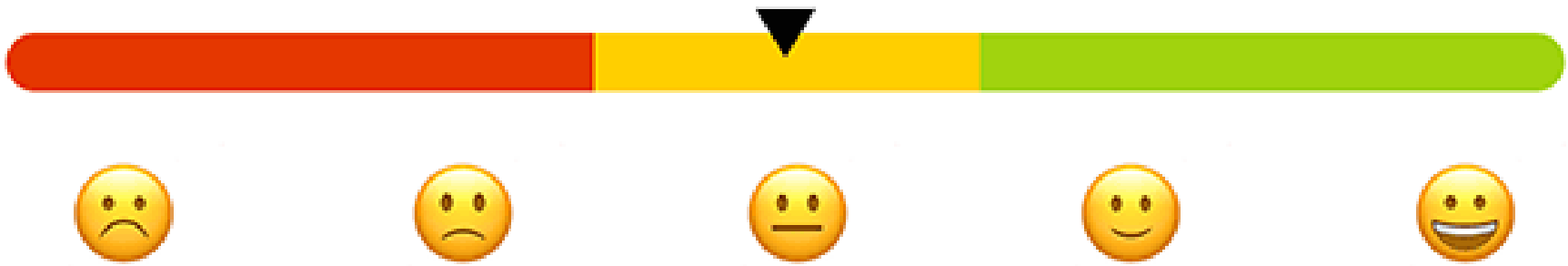
- Evolution of business decisions
- Need to process natural language
- **Applications of NLP**
- Process of natural language comprehension
- Pre-processing techniques for natural language
- Sentiment Analysis and Transfer Learning
- Annexure

We already know these



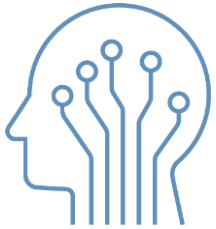


We understand these



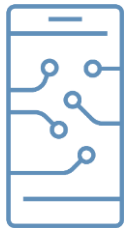
We already know these

NLP finds applications across businesses

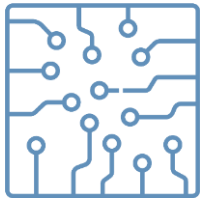


Sentiment analysis

Market intelligence



Chatbots



Intelligent call routing

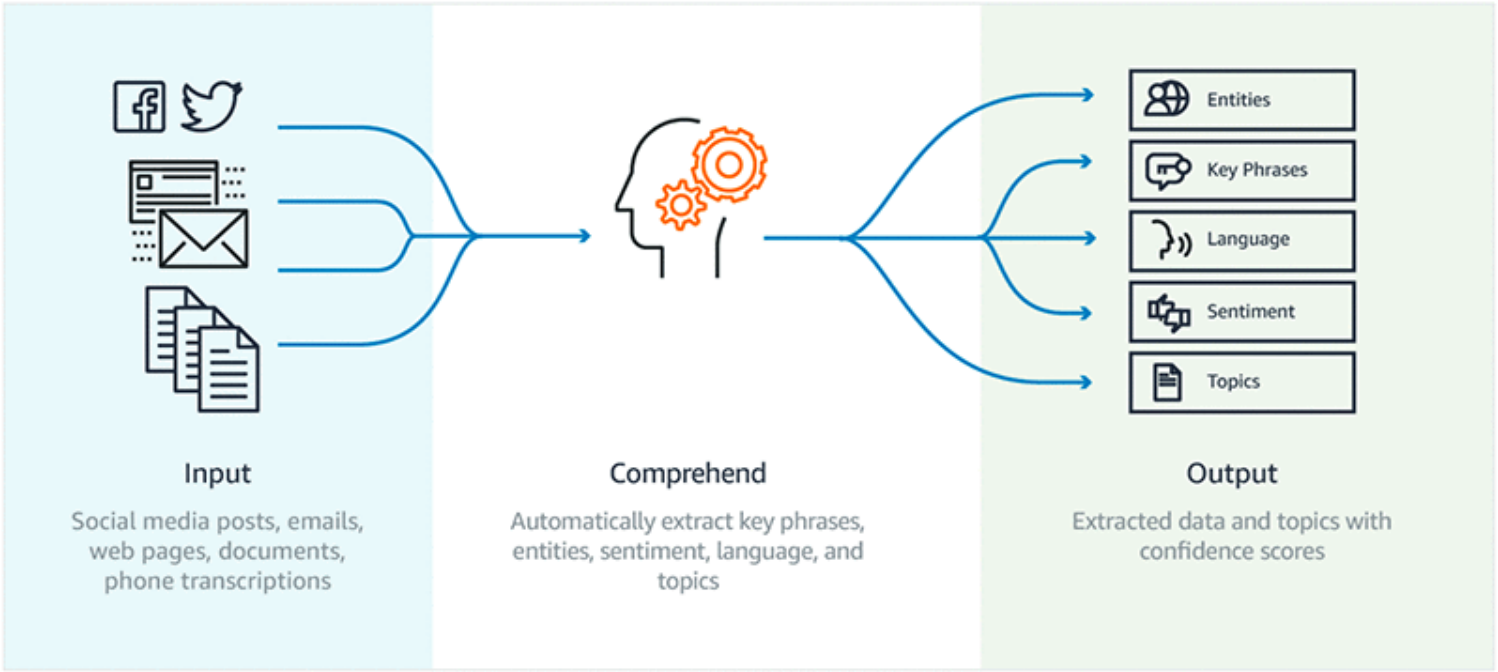
Customer service

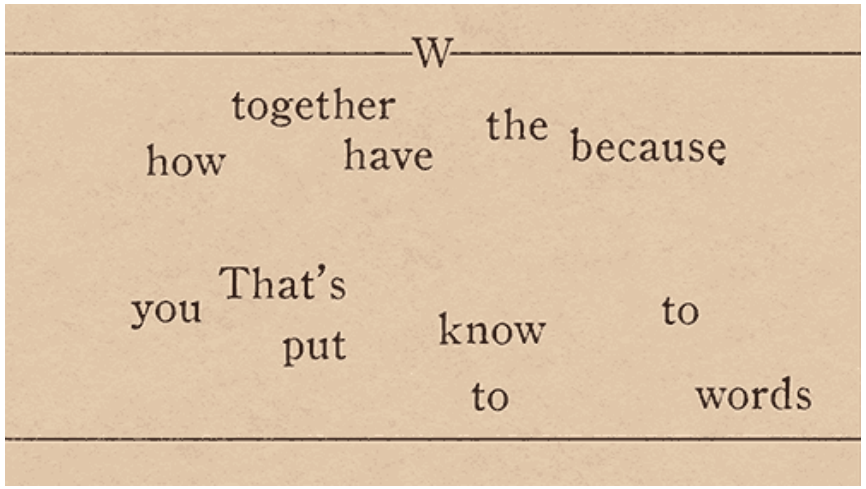
Intelligent reporting

Agenda

- Evolution of business decisions
- Need to process natural language
- Applications of NLP
- **Process of natural language comprehension**
- Pre-processing techniques for natural language
- Sentiment Analysis and Transfer Learning
- Annexure

Natural Language is interpreted intuitively by humans but computers require an algorithm to comprehend the data





Capabilities of NLP Span across syntax, semantics, discourse, and speech tasks

- Text classification : Web searching, information extraction, language identification, sentiment analysis
    - Name entity recognition (NER) : Classify entities (Microsoft, Picaso etc.) into organizations, people, locations, time, dates and so on
    - Part of speech tagging : tags nouns, verbs, phrasal verbs and so on. which can be used for text to speech, information extraction etc.

Capabilities of NLP Span across syntax, semantics, discourse, and speech tasks

- Semantic phrasing and question parsing : automatically answering questions asked in natural languages including definition questions, biographical questions, multilingual questions, etc.
- Paraphrase detection : Identifies if natural language questions are equivalent. This is used in information extraction
- Language generation and summarization
- Character recognition
- Spell check

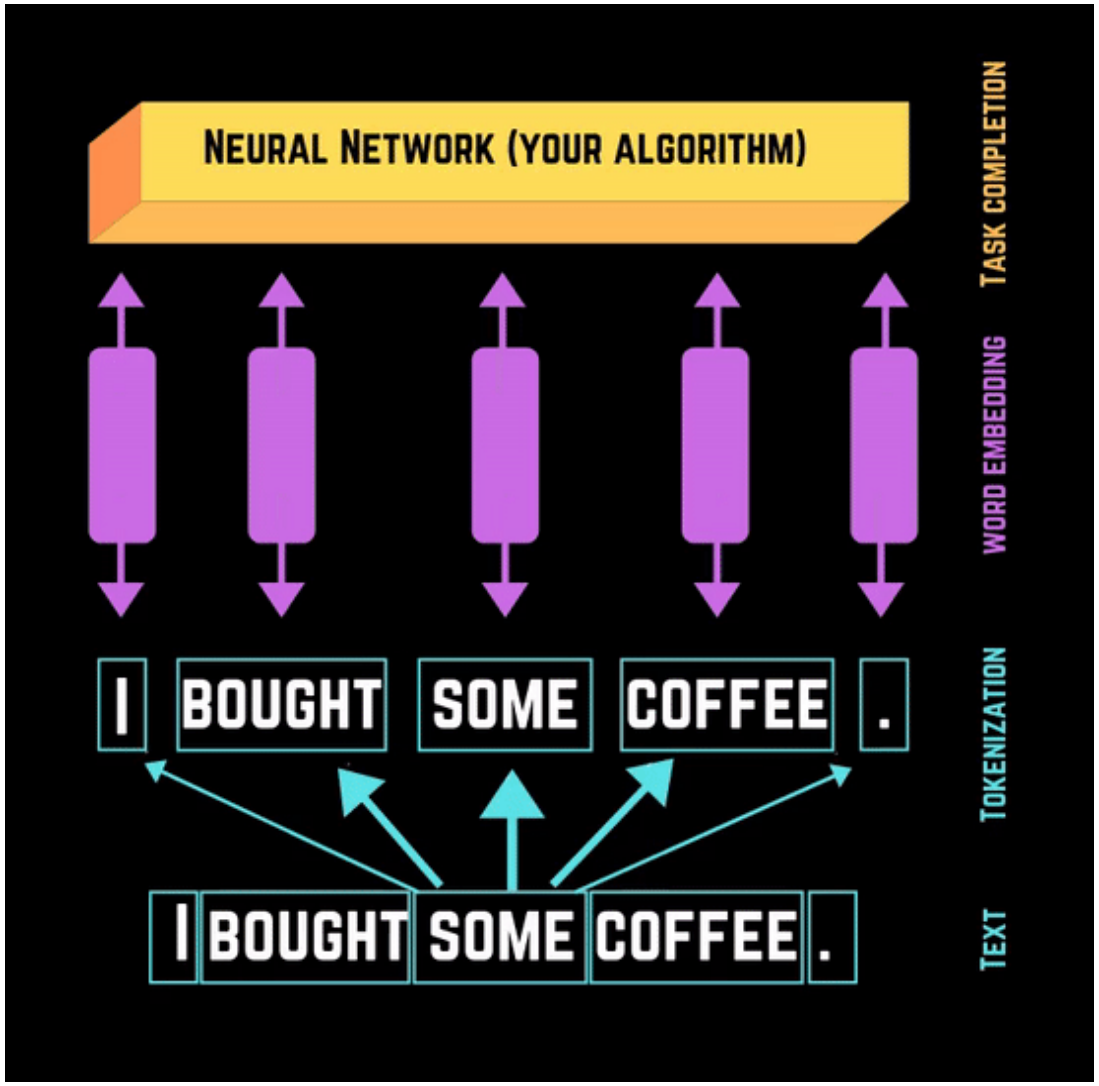
Agenda

- Evolution of business decisions
- Need to process natural language
- Applications of NLP
- Applications of NLP
- **Pre-processing techniques for natural language comprehension**
- Sentiment Analysis and Transfer Learning
- Annexure

Frequently used pre-processing techniques in NLP

- Bag of words
- Tokenization
- Stop word removal
- Spell-check
- Stemming
- Lemmatization
- Topic Modeling

Tokenization is the process of segmenting running text into tokens of words based on spaces

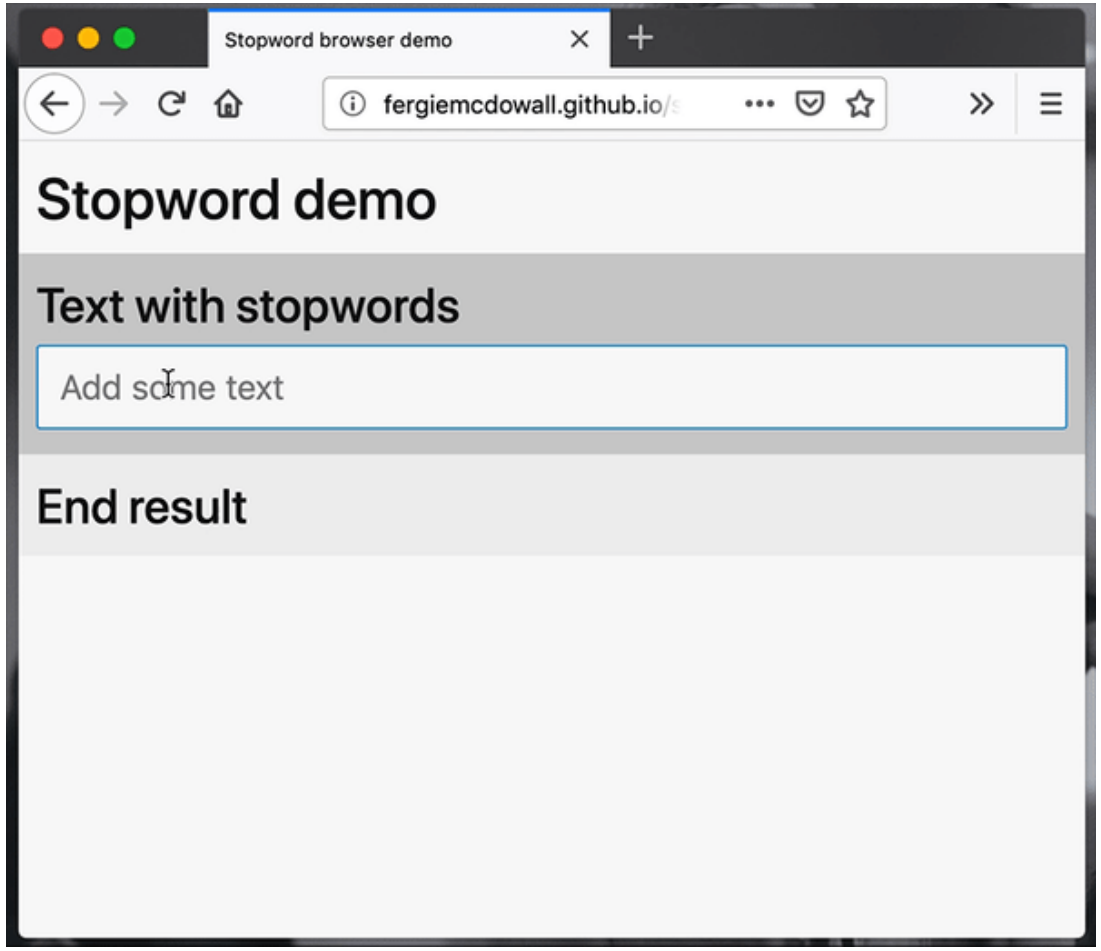


Words are flowing out like endless rain

Words	are	flowing	out	like	endless	rain

- Removes special characters and spaces based on level of tokenization (Low Level and High Level)
- Problematic for space separated words like San Francisco or New York, abbreviated words like dr. and biomedical and mathematical data because of hiphens and paranthesis

Stop Words Removal includes getting rid of common language articles, pronouns and prepositions such as “and”, “the” or “to” in English



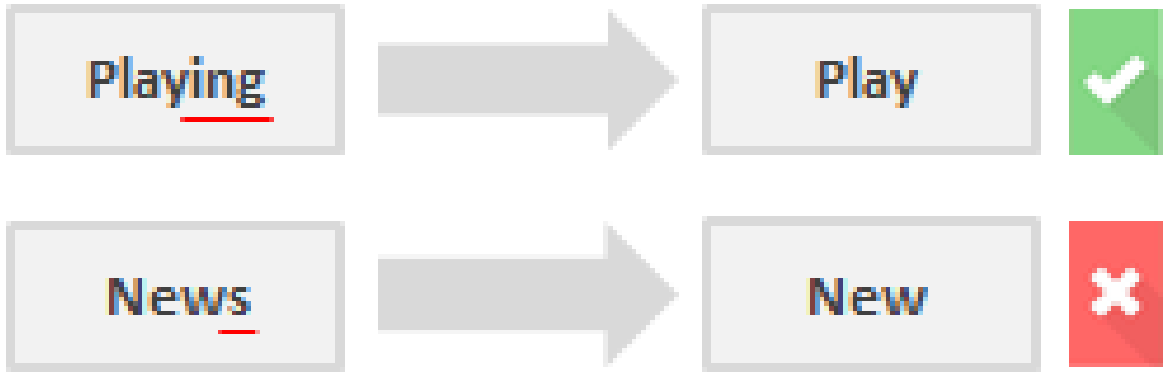
The apple was red in colour

apple, red, colour

- A stop word is a commonly used word (such as “the”, “a”, “an”, “in”) that a search engine has been programmed to ignore, both when indexing entries for searching and when retrieving them as the result of a search query.

Stemming is the process of slicing the end or the beginning of words with the intention of removing affixes (prefixes, suffixes)

> *Playing (‘play’ ‘ing’), Plays (‘play’ ‘s’)*
>
> *Guitarist (‘guitar’, ‘ist’)*



- Python and R languages have different libraries containing affixes and methods
- Stemming is also used to spell check words
- It is a fast algorithm as it is based on string operations

Lemmatization is more effective & computationally intensive them stemming

Lemmatization extracts root word from different forms (tense, degree of comparision etc.)

Better, best, good -> good

Go, went, gone -> go

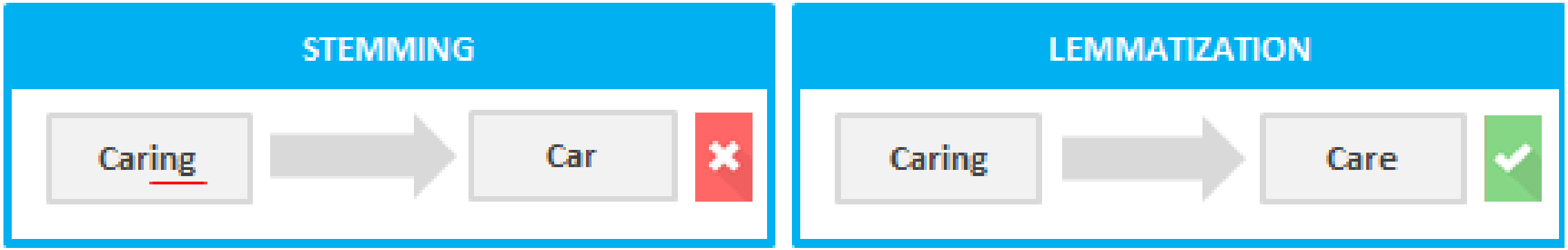
Running, runs, ran -> run

Lemmatization can be used with parts of speech to make it more accurate

Bat :

verb -> activity of batting

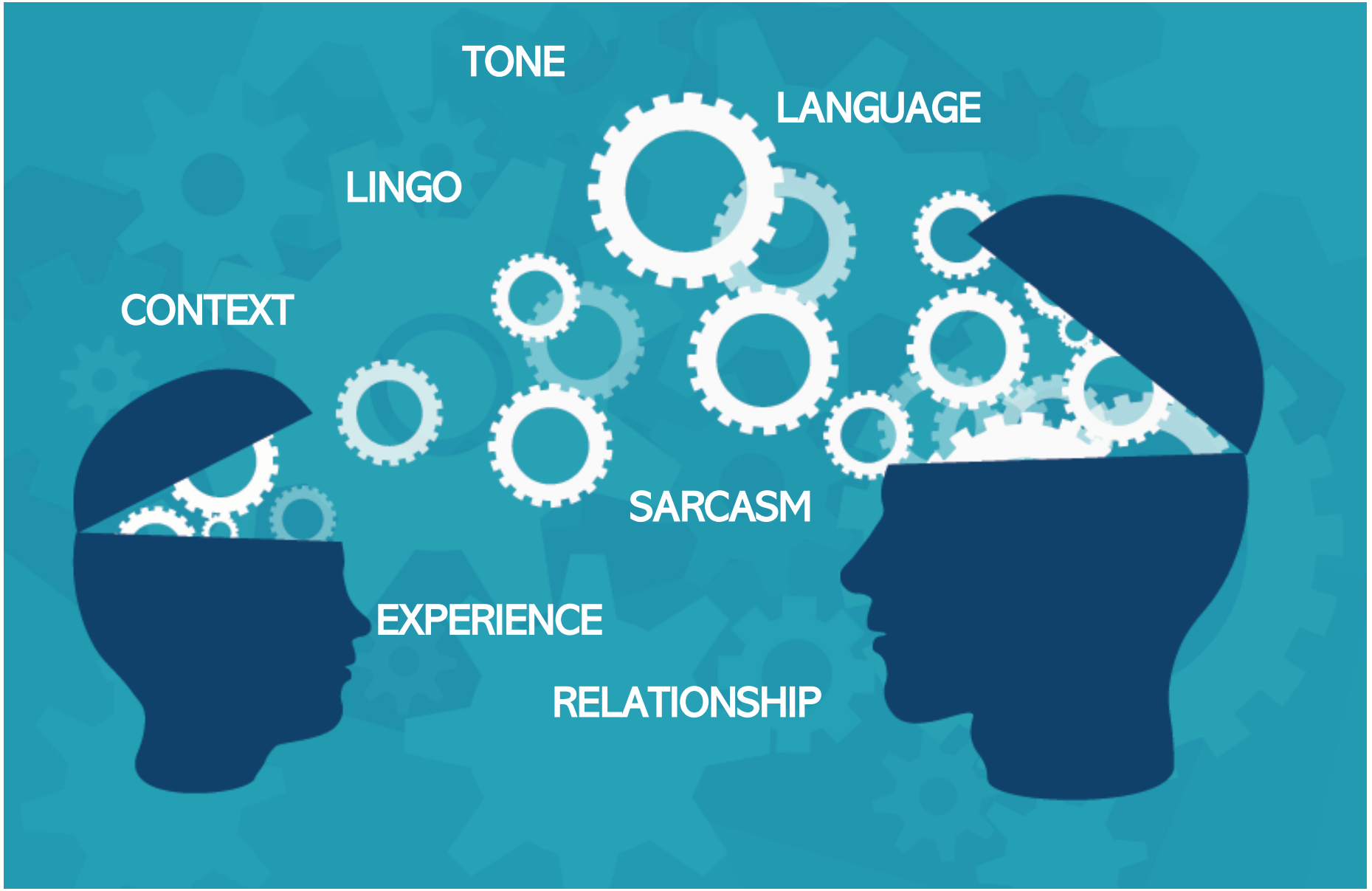
noun -> animal or the object



Agenda

- Evolution of business decisions
- Need to process natural language
- Applications of NLP
- Applications of NLP
- Pre-processing techniques for natural language
- **Sentiment Analysis and Transfer Learning**
- Annexure

How do humans perceive sentiment?



Building a sentiment comprehension system from scratch is not the most effective approach

Requires large amounts of data, Processing power & Time to be accurate

- Frameworks – PyTorch, Keras, Tensor Flow etc require large amounts of data
- High Data Requirement & resource intensive to build deep learning models from scratch

How does Transfer Learning simplify the process?

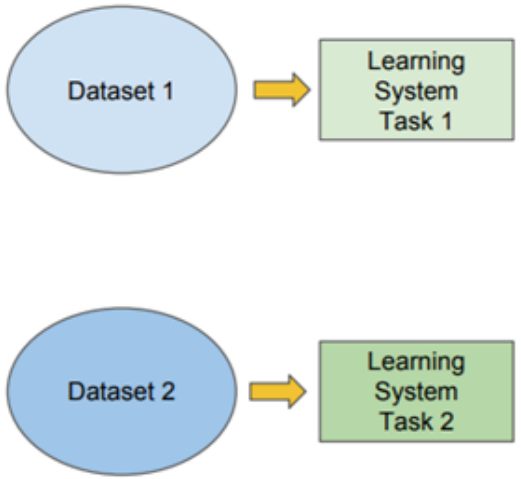
Transfers knowledge from one Domain or Task to Another and helps retain several features -

- Typically used when there is less data for Task 2
- Task 1 and Task 2 are similar and utilise similar features

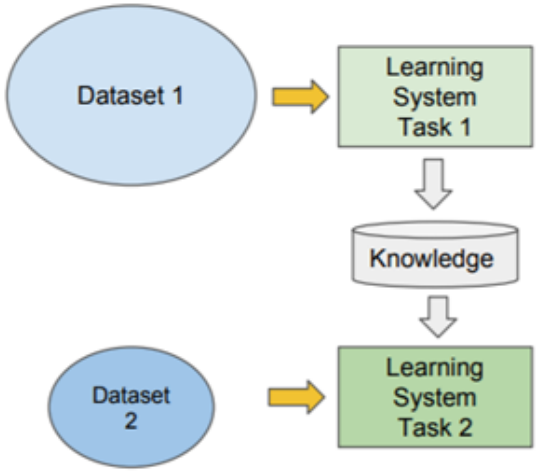
Works very well for unstructured data because many features can be extracted

Traditional ML vs Transfer Learning

- Isolated, single task learning:
  - Knowledge is not retained or accumulated. Learning is performed w.o. considering past learned knowledge in other tasks



- Learning of a new tasks relies on the previous learned tasks:
  - Learning process can be faster, more accurate and/or need less training data



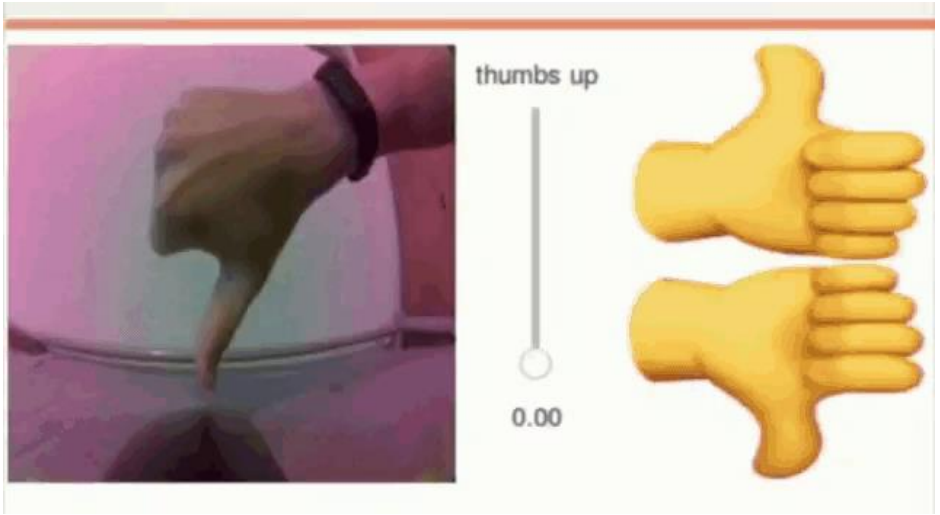
Example:

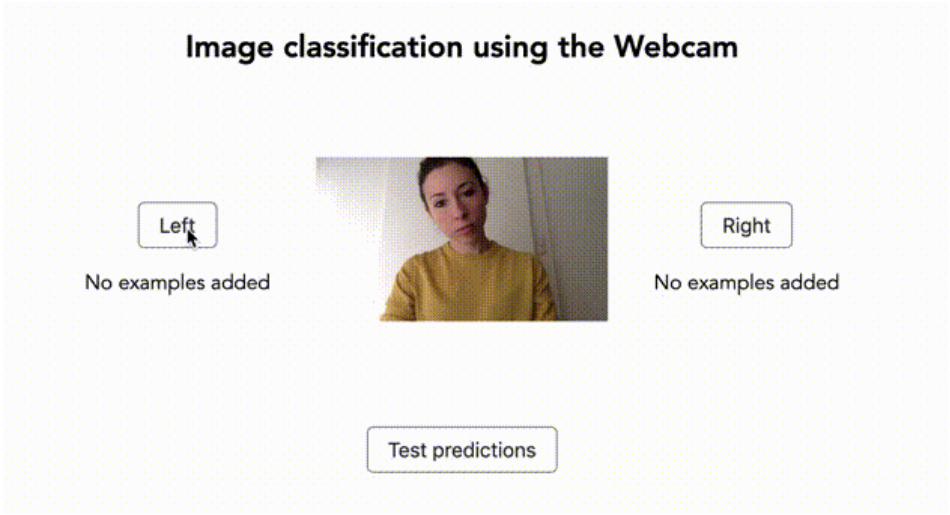
Task 1 - Generic object detection of 100 object classes

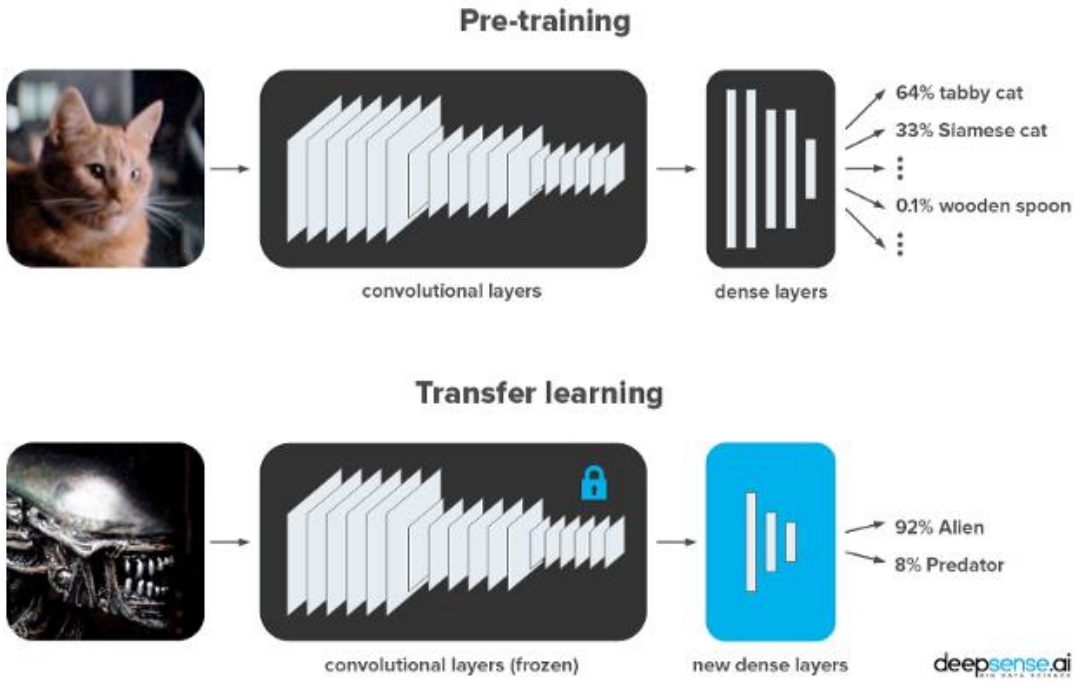
Task 2 - Detection of different types of bags

- Use 'bag' class
- define 3 subclasses 'bag pack', 'handbag', 'suitcase'

Transfer Learning Examples

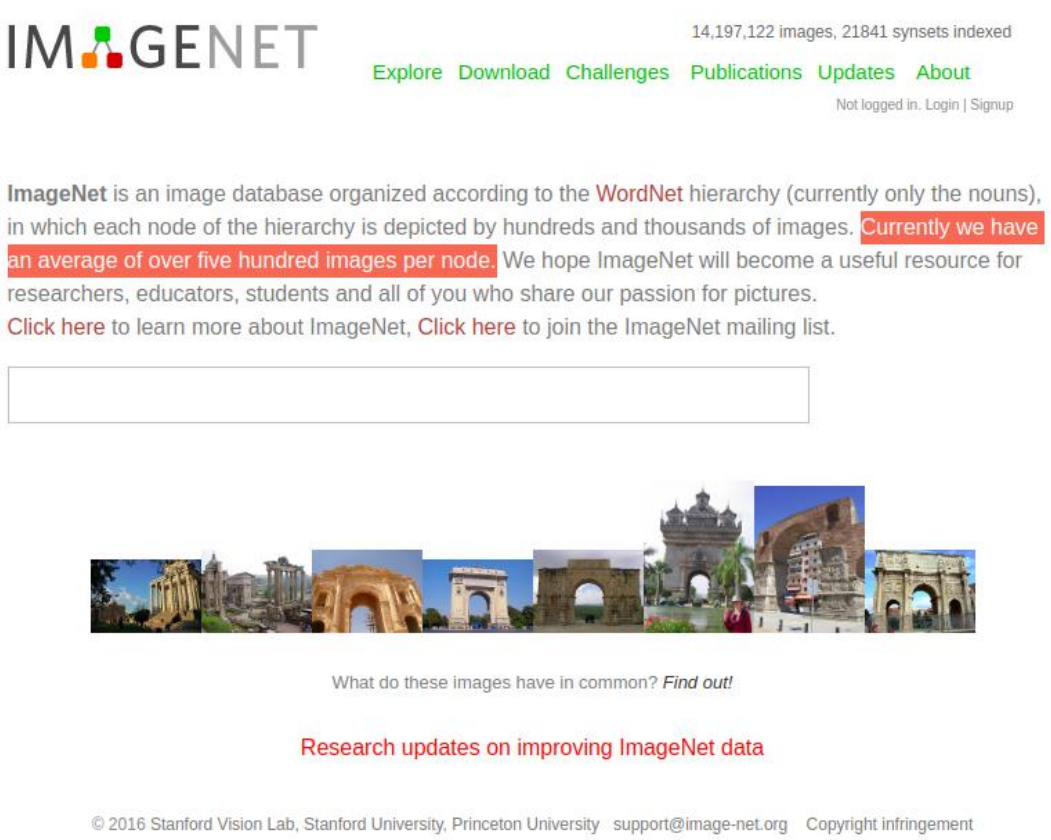


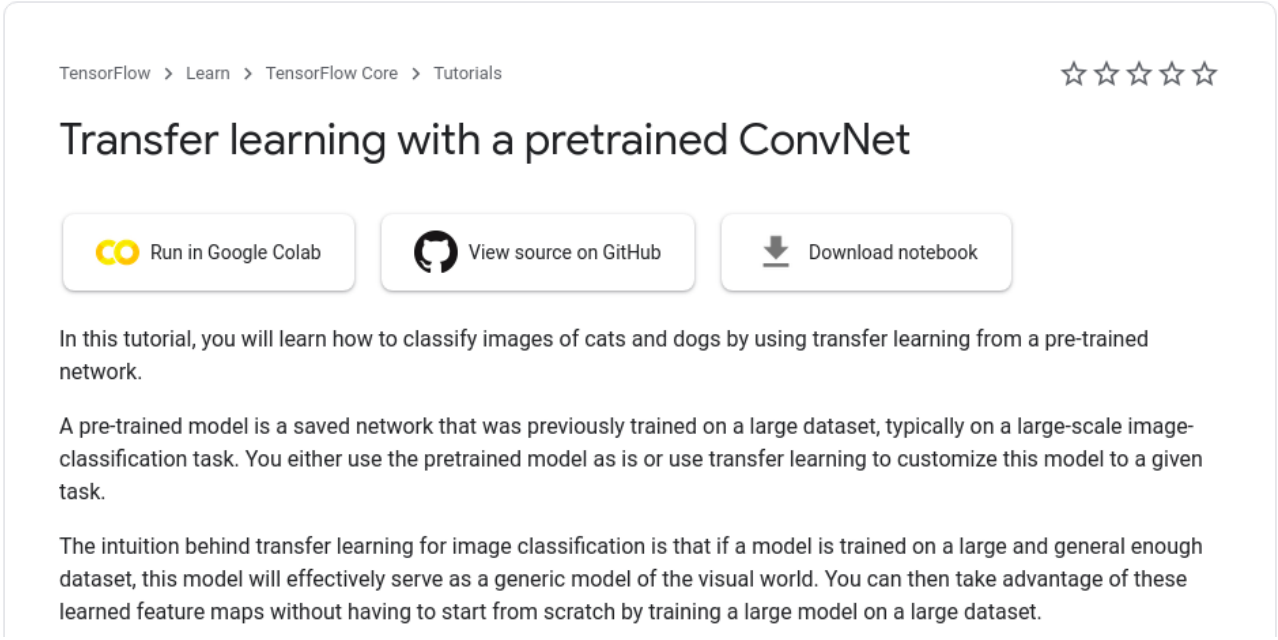




Pre-trained Deep-Learning Models

- Trained by Industry experts and Institutes with extensive research by Mathematicians, Scientists etc.
- Made with industry agnostic weights – minimum biases
- Time saving and universal structure





Agenda

- Evolution of business decisions
- Need to process natural language
- Applications of NLP
- Applications of NLP
- Pre-processing techniques for natural language
- Sentiment Analysis and Transfer Learning
- **Annexure**

Bag of words is a commonly used model that allows you to count all words in a piece of text by creating an occurrence matrix

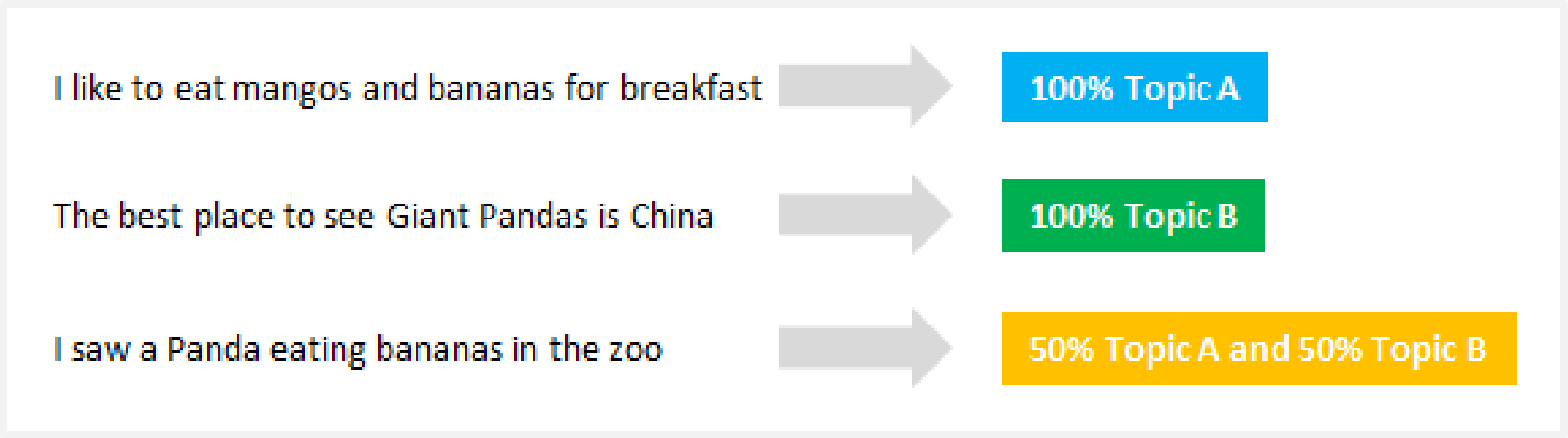
> *Words are flowing out like endless rain into a paper cup,*
>
> *They slither while they pass, they slip away across the universe*

	words	rain	a	paper	they	slip	the	universe	...
Words are flowing out like endless rain into a paper cup,	1	1	1	1	0	0	0	0	...
They slither while they pass, they slip away across the universe	0	0	0	0	3	1	1	1	...

- Downside is the absence of semantic meaning and context and stop words add noisy data
- The Term Frequency Inverse Document Frequency (TFIDF) can be created to compensate for the word repetition `

Topic Modelling is as a method to uncover hidden structures texts or documents by clustering topics based on their contents

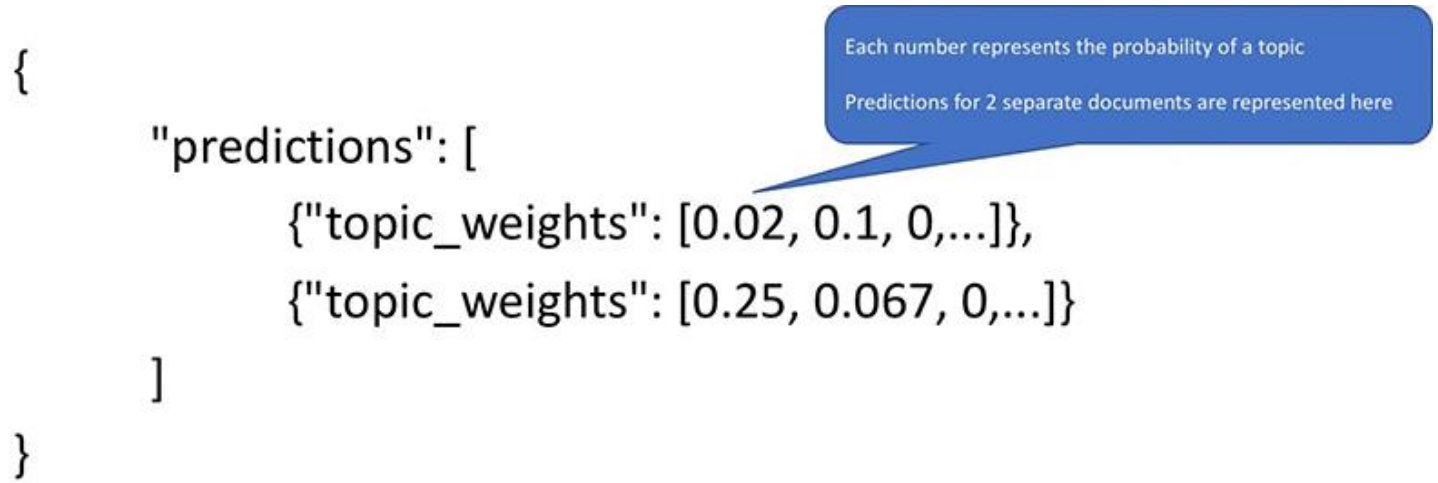
> *An unsupervised learning method called Latent Dirichlet Allocation (LDA) is most commonly used*



- Topic modelling is extremely useful for classifying texts, building recommender systems (e.g. to recommend you books based on your past readings)

Topic Modelling is as a method to uncover hidden structures texts or documents by clustering topics based on their contents

The user defines the number of topics (n) they want to extract from a text. LDA then assigns words to the most relevant topic by taking into account the probability that it belongs to that topic. This is an iterative process till the best topic clusters are obtained



Appendix

- Sentiment Analysis in Business (watch Nike's ad with Kolin Kaepernick)
- Applications of ANN in NLP
- Applications of NLP in business
- Information Extraction using Parts of Speech tagging and Entity Extraction
- IMPORTANT: Power of Hash tagging (in context of sentiment analysis)
- What is Transfer Learning? Exploring the popular Deep Learning Approach

Thank you