



# NimbleBot

Terone & Thumb

Growth Session #27 - August 22-23 2019

- Gives accessibility to newcomers and veterans alike to our knowledge base.
- Allows us to improve our understanding with machine learning.
- We're trying out a new approach on application of machine learning.
- Commonly used in images and raw number data, however, we're testing same principles with text.

## Goals

- Analyze methods of preprocessing text to vectors suitable for machine learning.
  - Mapping Questions to Answers.
  - Converting text to multi-dimensional arrays.
- Building the chatbot and integrating to Slack.

## Achievement and Progress

- Tokenized our corpus and stemmed the data.
- Encoded our labels with indexing to the topic.
- Associated our encoded labels with our multidimensional array.

## The equation of ...

$$\begin{aligned}(f * g_N)[n] &= \sum_{m=0}^{N-1} f[m]g_N[n - m] \\&= \sum_{m=0}^n f[m]g[n - m] + \sum_{m=n+1}^{N-1} f[m]g[N + n - m] \\&= \sum_{m=0}^{N-1} f[m]g[(n - m) \bmod N] \triangleq (f *_N g)[n]\end{aligned}$$

'Don't see the play because I haven't seen the play'

Tokenizer

['do', 'n't', 'see', 'the', 'play', 'because', 'i', 'have', 'n't', 'seen', 'the', 'play']

Indexing

[10, 3, 29, 8, 51, 16, 7, 5, 3, 23, 8, 51]

# Tokenization

$V$  = Vocabulary size

$V_{th}$  value

the = [ 1 , 0 , 0 , 0 , 0 , ... , 0 ]

a = [ 0 , 1 , 0 , 0 , 0 , ... , 0 ]

he = [ 0 , 0 , 1 , 0 , 0 , ... , 0 ]

she = [ 0 , 0 , 0 , 1 , 0 , ... , 0 ]

go = [ 0 , 0 , 0 , 0 , 1 , ... , 0 ]

...

word  $V$  = [ 0 , 0 , 0 , 0 , 0 , ... , 1 ]

# Array Mapping



```
Numo::Int32#shape=[10]  
[6, 4, 5, 3, 8, 7, 9, 0, 1, 2]  
["Topic", "Company Ideology Mission", "Company Ideology Vision", "Company Ideology", "Values Core-  
Values", "Values Aspirational-Values", "Zode Conventions Formatting", "Code Conventions Naming-Files",  
"Code Conventions Naming-ID-Class-Names", "Code Conventions Tables"]
```



# Array Mapping



## Question

What is the company mission? What is the company goals? What is the goals of the company? What is the mission of the company?

What is the company vision? What is the future of the company? What does the company want to be?

What is the company ideology?

What does the company value?

What should we aspire to be?

What are the formatting conventions for web?

How do you name your files? What is the company's naming convention for naming files?

How do you name your class names and IDs? What is the company's naming convention for naming classes?

What is the naming convention for tables?

# Array Mapping

```
{ "goal"=>0, "mission"=>1, "vision?"=>2, "goals?"=>3, "company?"=>4, "futur"=>5, "want"=>6, "be?"=>7,
"ideology?"=>8, "value?"=>9, "aspir"=>10, "format"=>11, "convent"=>12, "web?"=>13, "files?"=>14,
"company'"=>15, "class"=>16, "ids?"=>17, "classes?"=>18, "question"=>19, "tables?"=>20, "compani"=>21,
"mission?"=>22}
```

```
Numo::Int32#shape=[10,23]
```

```
[[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0],
 [1, 1, 0, 1, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 1],
 [0, 0, 1, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0],
 [0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0],
 [0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0],
 [0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
 [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0],
 [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 2, 1, 0, 0, 0, 0],
 [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 1, 1, 1, 0],
 [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0]]
```



Stay tuned!

# Thanks!

## Contact Nimble

[nimblehq.co](https://nimblehq.co)

[hello@nimblehq.co](mailto:hello@nimblehq.co)

## Bangkok

399 Interchange 21 Sukhumvit Road, Unit  
#2402-03, Klong Toei, Wattana, Bangkok  
10110, Thailand

## Singapore

28C Stanley St, Singapore 068737

## Hong Kong

20th Floor, Central Tower  
28 Queen's Road, Central, Hong Kong

