

# Poco: Text Classification With Zero Shot and Few Shot Learning

19.11.2022 (Sat) 1pm-2pm



# **Speaker Profile**



#### **Phyo Thu Htet**

Founder and CEO at Simbolo

Introduction to Al Class Instructor at Simbolo

#### **Experiences Regarding AI**

- Al Engineer and Consultant at Koe Koe Tech
- Al Team Executive (Academic Department, Thate Pan Hub)
- NLP Researcher/ Research Fellow, Myanmar NLP Lab



#### **Outline**

- Text Classification
- Tokenization: Preprocessing
- Tokenization Demo Session
- Traditional Supervised Learning
- Text Classification With Zero Shot
- Text Classification With Zero Shot Demo Session (Hugging Face)
- Text Classification With Siamese Network Few Shot





# What is Text Classification? (Discussion)





# **Text Classification**



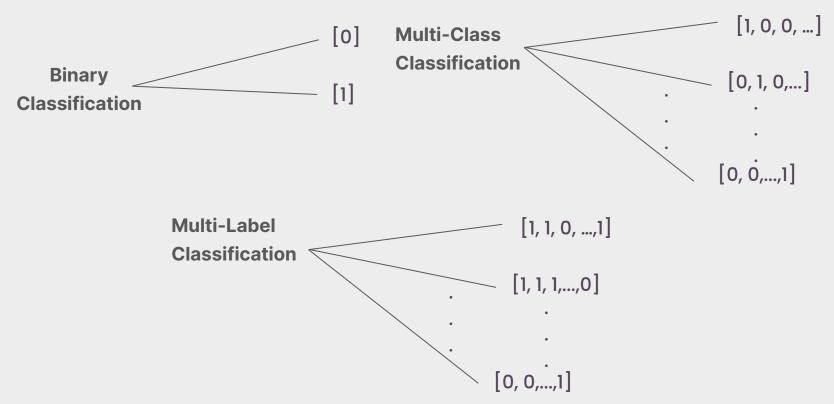
#### **Text Classification**

- **Text Classification**: Text classification is the technique of NLP used to categorize texts into different groups.
- Text Classification is one of the Natural Language Understanding (NLU) Tasks.
- The examples include Email-Spam Detection, Hate Speech Detection and Sentiment Analysis.





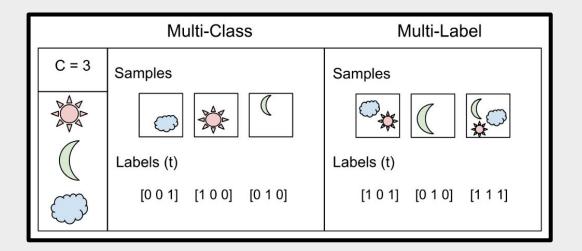
# **Types of Classification**



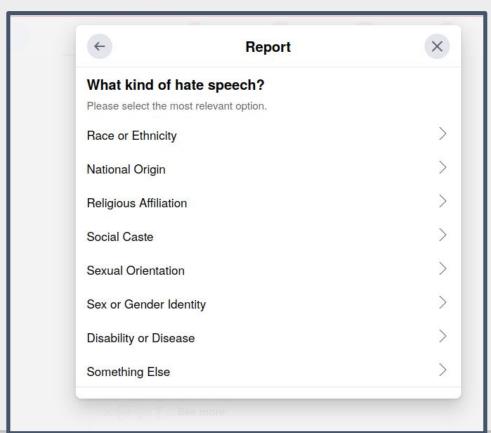


- Generic Binary Classification: f:text → {0,1}<sup>M</sup> i.e., f(t) = P(y<sub>i</sub>|t) ∀ i ∈ {1...M}
- For multi-class problems the labels are mutually exclusive i.e., only one of them can be true.













# **Tokenization: Text Preprocessing**



# **Some Preprocessing Technique**

- 1. Tokenization (တူညီသော အပိုင်းလေးများ ဖြတ်ပေးခြင်း)
- 2. Text to Number (စာမှ နံပါတ်သို့ ပြောင်းပေးခြင်း)
- 3. Padding (စာကြောင်းအလျားများတူညီအောင် ပြုလုပ်ခြင်း)
- 4. Other





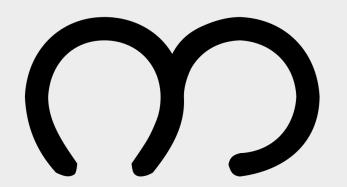
#### **Tokenization**

Tokenization is a way of separating a piece of text into smaller units called tokens.



#### What is a Character?

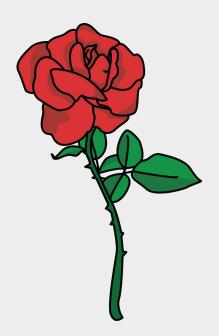
A *character* is a typographical symbol. For example, any of these could be classified as characters: က ခ ာ ု ် ွ ျ ၁ ၂ ၃ ။







# What is Syllable?



Syllable is a unit of pronunciation forming the whole or a part of a word; for example, there are two syllables in လှတယ်, and three syllables in နှင်းဆီပန်း.



#### **Character Tokenization**

Input: အမုန်းမပွားရဝူးနော်

Output: အမ ုန ် း မ ပ ွ ာ း ရ ဝ ူ း န ေ ာ ်

#### Syllable Tokenization

Input: အမုန်းမပွားရဝူးနော်

Output: အ မုန်း မ ပွား ရ ဝူး နော်



# Multilingual Semi-Syllable Break

Semi-syllable-break for Lao, Kannada, Oriya, Gujarati, Malayalam, Khmer, Telugu, Bengali, Sinhala, Tamil, Shan, Mon, Pali and Sanskrit, Sgaw Karen, Western Poh Karen, Eastern Poh Karen, Geba Karen, Kayah, Rumai Palaung, Khamti Shan, Aiton and Phake, Burmese (Myanmar), Paoh, Rakhine Languages, Word break for English, Char break for other language



- Here, semi-syllable does not refer to the minor syllable in phonology. Instead, it
  is new tokenization that does not break into a full syllable mode. Now I found
  that it is useful in keyword detection to reduce False Positive errors. (I may
  explain why keywords detection later)
- The beauty of this tokenization would be you don't need to know much about the nature of the specific language. It will especially work for a similar script like Brahmic Script.



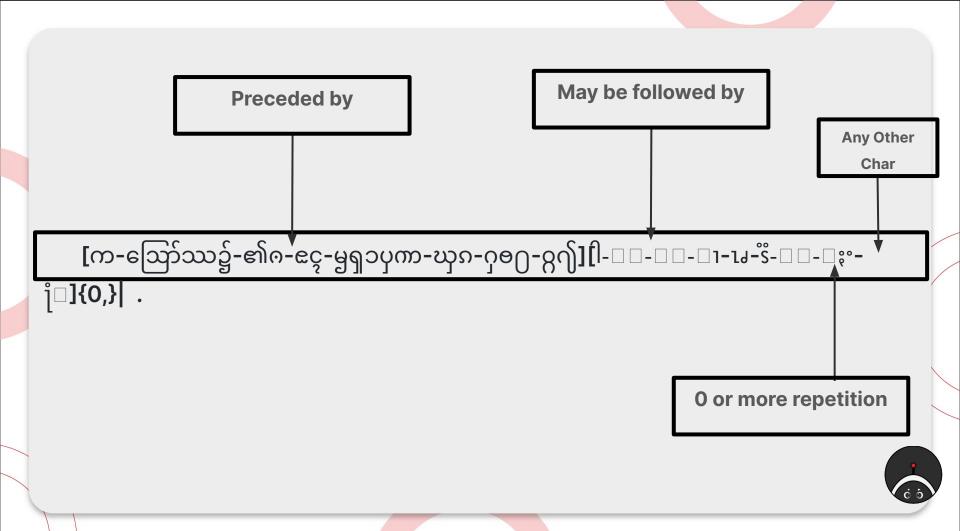
- Input: ဝါဆိုဝါခေါင် ရေတွေကြီးလို့ သပြေသီးမှည့် ကောက်စို့ကွယ် Output: ဝါ....ဆို....ဝါ....ခေါ....င်.....ရေ....တွေ....ကြီး....လို့.... သ....ပြေ.... သီး....မှ....ည့်.... ကော....က်....စို့....ကွ....ယ်....





Symbol	Meaning	Example
•	any character	a.c matches abc
	character . itself is represented as \.	
*	zero or more repetition of the previous character	a* matches aaa
+	one or more repetition of the previous character	a+ = aa*
?	zero or one repetition of the previous character	a? matches a or null character
	or connective	a b matches a or b
[]	any character within the bracket	[abc] = a b c
\a	any alphabetic characters	\a = [abcz]
\d	any numeric characters	\d = [0129]
[^abc]	any character other than a, b or c	[^abc] matches e



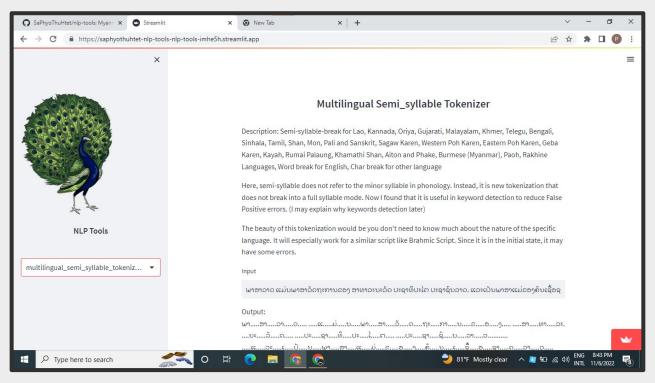




# **Tokenization: Demo Session**



### Demo



https://github.com/SaPhyoThuHtet/nlp-tools

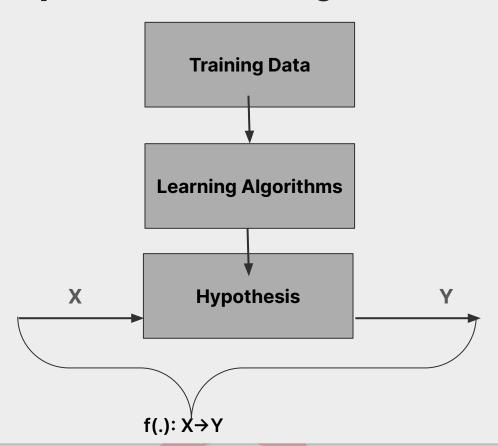




# Traditional Supervised Learning

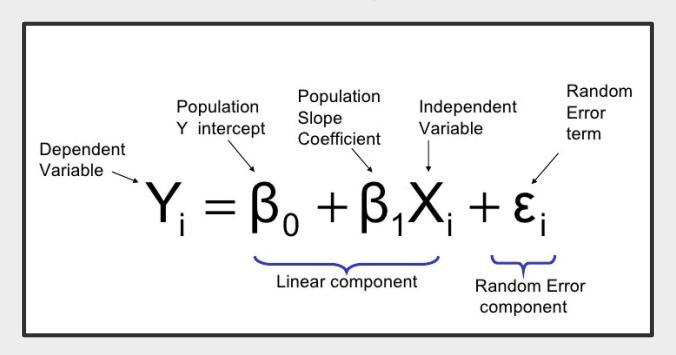


# Traditional Supervised Learning for Text Classification





# **Linear Regression**

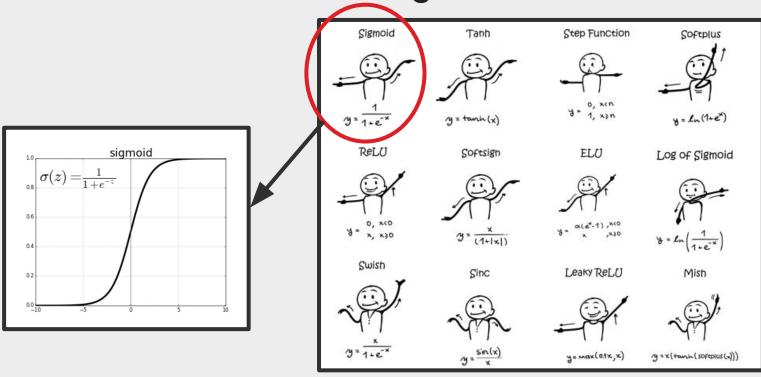


#### Ref:

https://towardsdatascience.com/how-are-logistic-regression-ordinary-least-squares-regression-related-1deab32d79f5



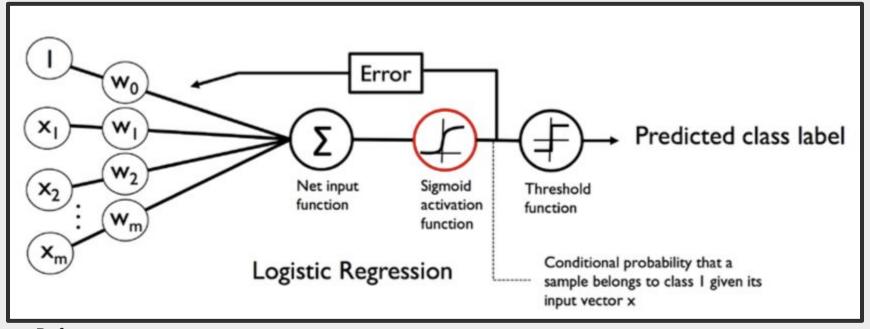
## **Add Sigmoid**



Ref: https://sefiks.com/tag/activation-function/



# Logistic Regression (Binary Classification Setting)



Ref:

https://www.researchgate.net/figure/Architecture-of-a-Logistic-Regression-Model-56\_fig 7\_334575492



$$f(X) = \begin{bmatrix} 0 & \text{if } P/t \\ 1 & \text{if } P/t \end{bmatrix}$$

- t = threshold
- p= probability score





# Do you think Supervised Learning require a lot of training data?





# Few Shots (Zero Shot, One Shot and Few Shot Intro)



# K-way and n-shot

k-way = the number of classes in the support set

n-shot = the number of samples of each class



#### n-shot

- Zero Shot Learning (ZSL)
- One Shot Learning (OSL) E.g. Face Recognition
- Few Shot Learning (FSL)

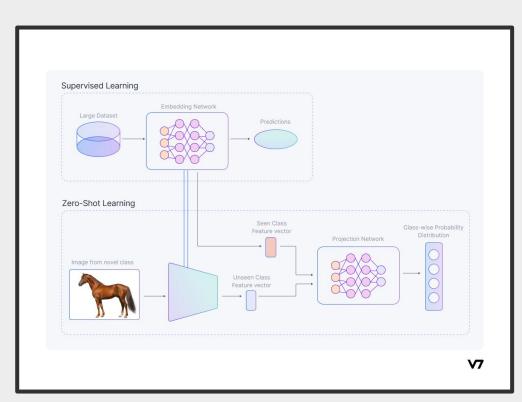




# **Zero Shot Learning**



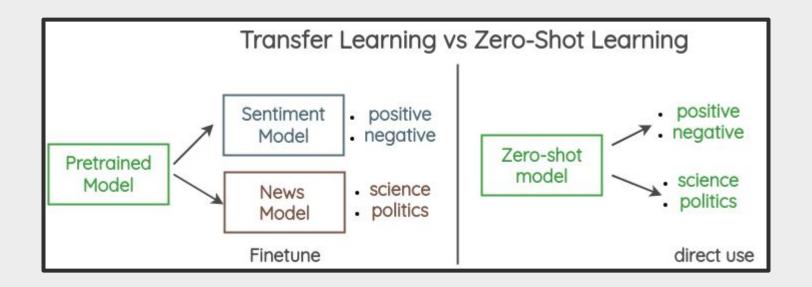
## **0-shot learning**



- Seen Classes
- Unseen Classes
- Auxiliary Information



Ref: https://www.v7labs.com/blog/zero-shot-learning-guide



Ref: Amit Chaudhary https://amitness.com/2020/05/zero-shot-text-classification/





With Zero Shot Learning, you need 0 training data.

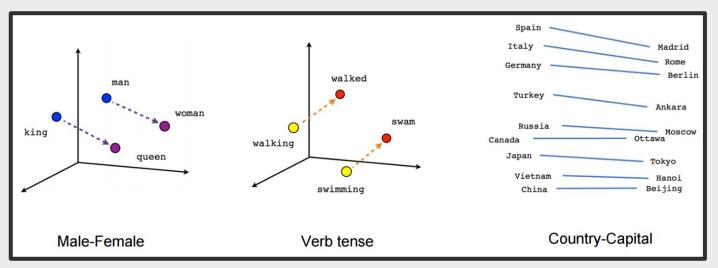


# Some approaches for Zero Shot Text Classification

- Latent Embedding
- Text Awareness Representation of Sentences (TARS)
- Natural Language Inference



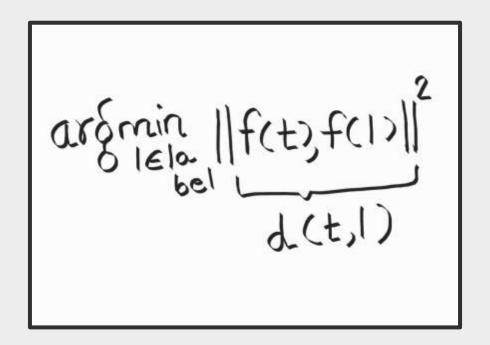
# **Embedding**



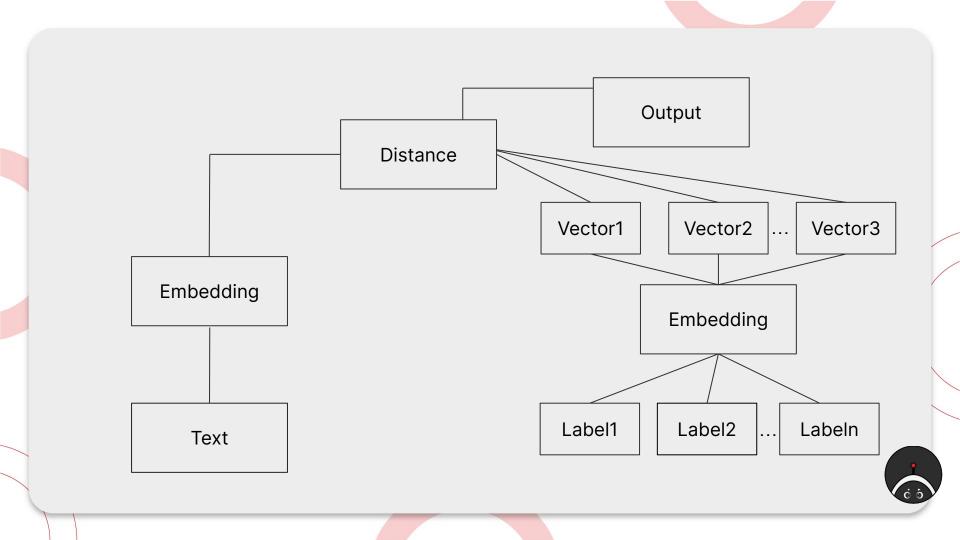
Ref: https://towardsdatascience.com/deep-learning-4-embedding-layers-f9a02d55ac12

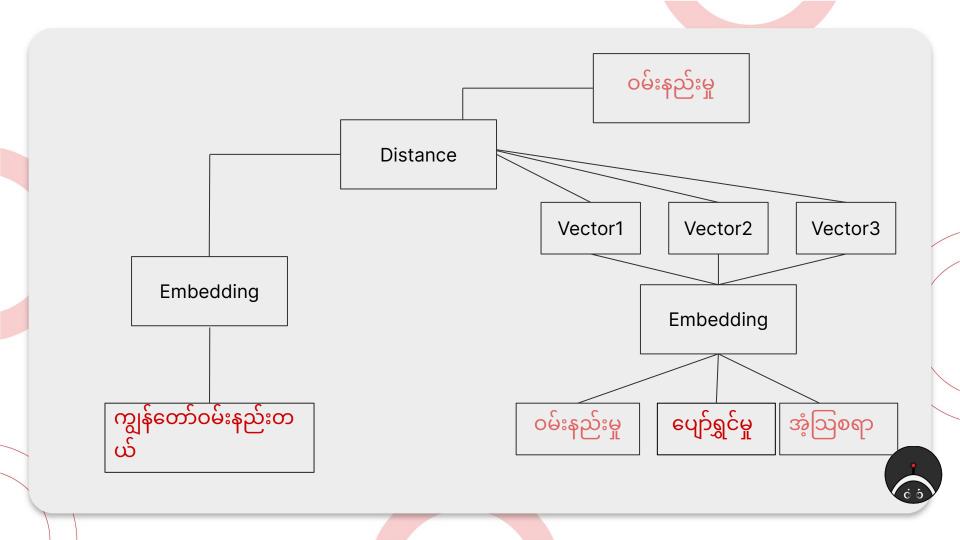


# **Zero Shot Text Classification with Latent Embedding**

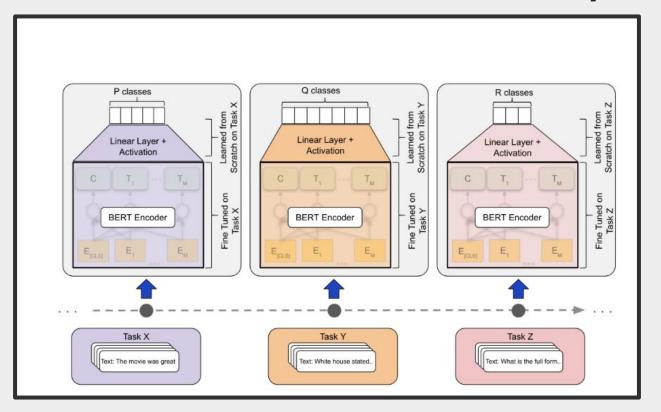






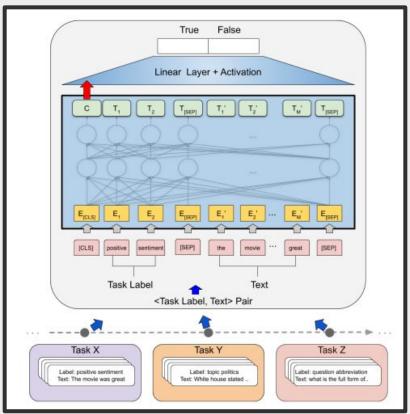


# **Traditional Text Classification Techniques**





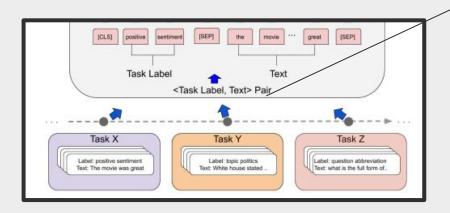
# **Text Awareness Representation of Sentences (TARS)**



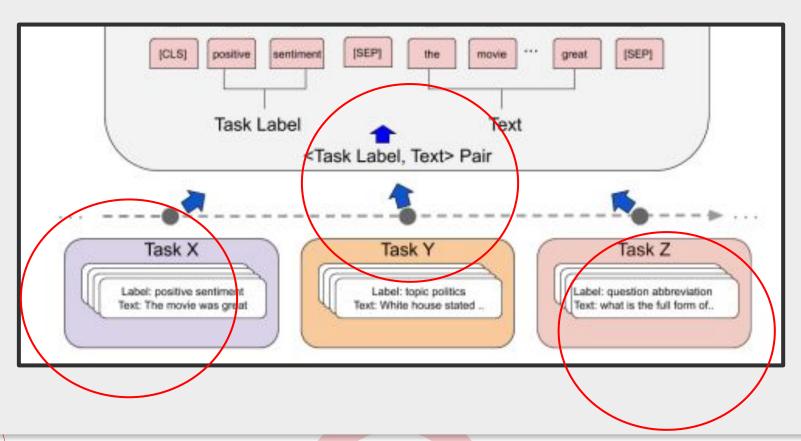


Ref: https://kishaloyhalder.github.io/pdfs/tars\_coling2020.pdf

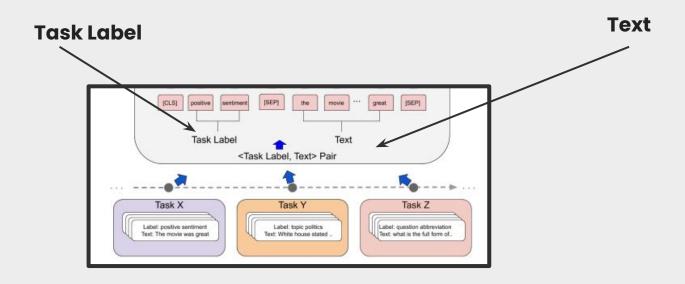
#### <Task Label, Text> Pair



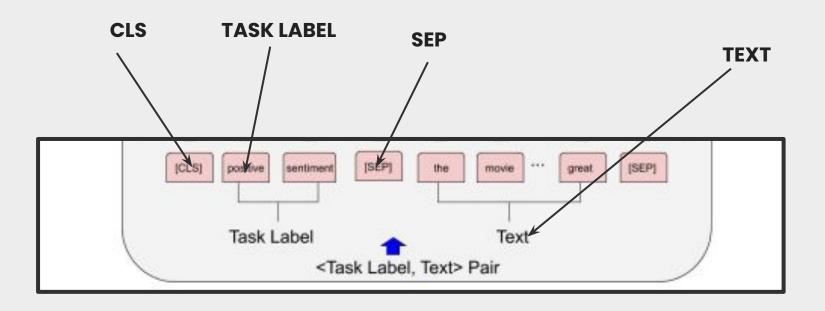




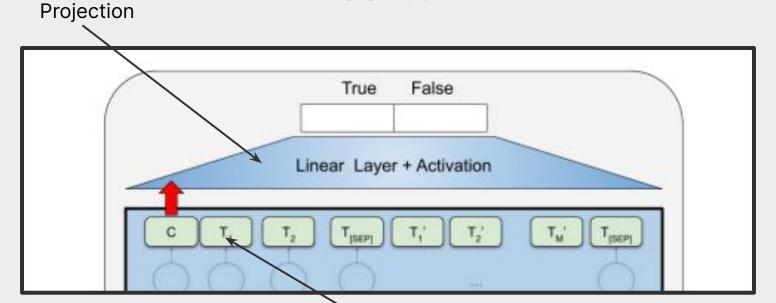








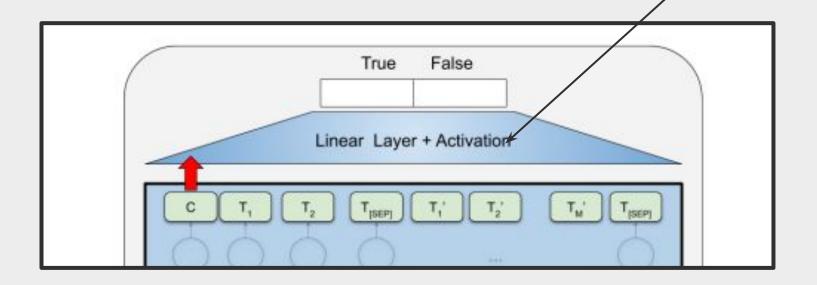




H-dimensional tensor produced by the encoder



$$\sigma(\vec{z})_i = rac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$





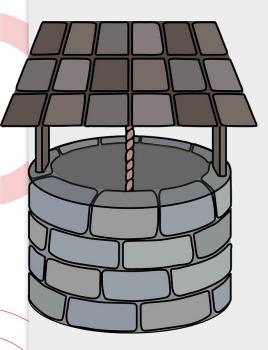
f:text 
$$\to \{0,1\}^M$$
 i.e.,  $f(t) = P(y_i|t) \ \forall \ i \in \{1...M\}$ 



f: <task label, text>  $\rightarrow$  {0,1} i.e. f(label(y<sub>i</sub>), t) = P(True | (y<sub>i</sub>,t))  $\forall$  i  $\in$  {1...M}



# Zero Shot Learning with Natural Language Inference



Premise: ခုရေတွင်းတူး ခု ရေကြည်သောက်လို့မရ။

- 1. The premise entails the hypothesis. Hypothesis: အလုပ်ကိစ္စ တစ်ခုကို ပြုလုပ်ရာတွင် ခုချက်ချင်းလုပ် ခုချက်ချင်းအောင်မြင်ရန် မဖြစ်နိုင်ပါ။
- 2. The premise contradicts the hypothesis. Hypothesis: အလုပ်ကိစ္စ တစ်ခုကို ပြုလုပ်တဲ့အခါ ခုချက်ချင်းလုပ် ခုချက်ချင်းအောင်မြင်မှာ ကျိန်းသေပဲ။
- 3. The premisse does not entail nor contradict the hypothesis.

Hypothesis: ရြတွင်းထဲကရေကြည်လေးသောက်ချင်လို့။

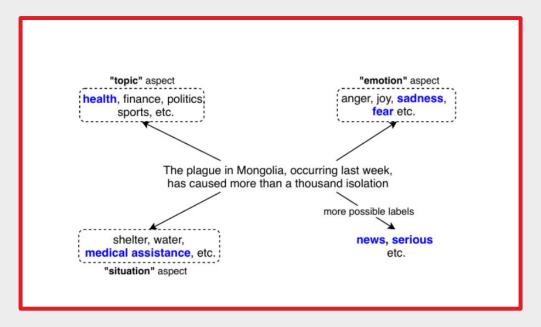


Premise	Label	Hypothesis
A man inspects the uniform of a figure in some East Asian country.	contradiction	The man is sleeping.
An older and younger man smiling.	neutral	Two men are smiling and laughing at the cats playing on the floor.
A soccer game with multiple males playing.	entailment	Some men are playing a sport.

 $Ref: http://nlpprogress.com/english/natural\_language\_inference.html\\$ 



# Zeroshot Text Classification (0-SHOT TC)



Ref: Yin et al., "Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach." arXiv preprint arXiv:1909.00161 (2019)



# These are important in zeroshot TC

```
sequence = "ကျွန်တော်ဝမ်းနည်းတယ်"
candidate_labels = ["ပျော်ရွှင်မှု", "အံ့ဩစရာ", "ဝမ်းနည်းမှု"]
hypothesis_template = "ဤစာသားသည် { \bar{b}} }."
```



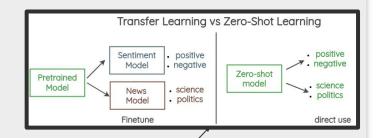
{'sequence': 'ကျွန်တော်ဝမ်းနည်းတယ်'

'labels': ['ဝမ်းနည်းမှု', 'ပျော်ရွှင်မှု', 'အံ့ဩစရာ'],

'scores': [0.8763300776481628, 0.07530076056718826,

0.04836917668581009]}





Supervised Learning	Zero Shot Learning
Tons of Annotated Training Data are needed.	Zero Training Data for your downstream tasks
You don't need to worry about the meaning/sentiment of the labels.	Candidates labels play a big role in zero shot learning and can be used as a hyperparameters.





# Zero Shot Learning (Demo)



#### gdg 2022

Zero Shot Example Notebook: https://colab.research.google.com/drive/1jocViLorbwWlkTXKwxCOV9HLTaDDgCaw?usp=sharing

https://github.com/SaPhyoThuHtet/nlp-tools/blob/main/README.md





# Few Shots With Siamese (Binary Classification Setting)



#### **Offensive**

- DNA အရ သူခိုးပိုးတွေအပြည့်နံ့တဲ့ ။ အဲဒီ နှစ် ယောက် ကို ဖင်လိုး ချင် တယ် တစ် ဆွေ လုံး ပေါ့ တစ်ဆွေလှုံး ကို ငါ လိုကောင်လီး ပဲ စိုက် မယ်
- 3. လိုးခိုင်းပလားပိုသက်ဝန်တာပေါ့
- 4. ဖာခေါင်းရှို့ တဲ့ နေရာ ဖာ သည် မရှို့ စမြဲ ပါ 5. မ ရှင်း ချင် ဘူး သာ ပြော တာ ဂယက် ထ လို့ ပျော် နေ တဲ့ မျက်နှာကြီး နဲ့ ဒါ မျိုးတေ့ က ရိုး နေ ပြီ
- ချတိုက် ကြ ကွာ သေ ကြ စမ်း
- ဖင် ထဲထည့် ထား လိုက်
- အေးပါကွာငပုရယ်
- အဆိုတော် လား ငါ့ လိုး ၍ သော သား ။ မုန်းစရာကောင်းတယ် အပြာတွေ တိုင်းပြည်အတွက်အမှိုက်တွေ အမှိုက်မှ ပြနါအသုံးပြုလို့ကို မရတဲ့အမှိုက်စုပ်တွေ

#### **Not Offensive**

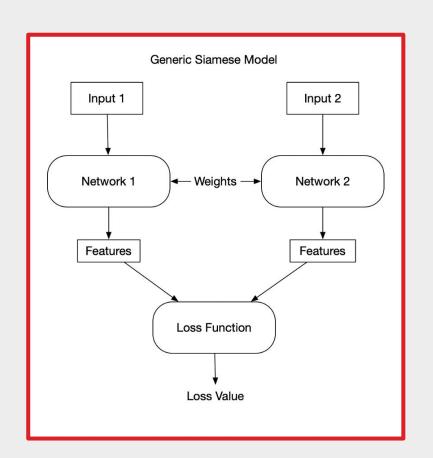
- 1. နေကောင်းလား။
- 2. ဒီနေ့ရာသီဥတု သာယာတယ်။
- 3. ကျန်းမာ ချမ်းသာကြပါစေ။
- 4. ပျော်ရွှင်ကြပါစေ။
- 5. GDG မှာ တွေ့ကြမယ်။
- 6. ORGANIZER ဘယ်သူလဲ?
- 7. Simbolo မှာ Al အတန်းတွေရှိတယ်။
- 8. အေးဆေးပဲ။
- 9. ပင်ပန်းရင် အနားယူပါ။
- 10. အားလုံး အဆင်ပြေကြပါစေ။

# Support Set (2 ways 10 shots)

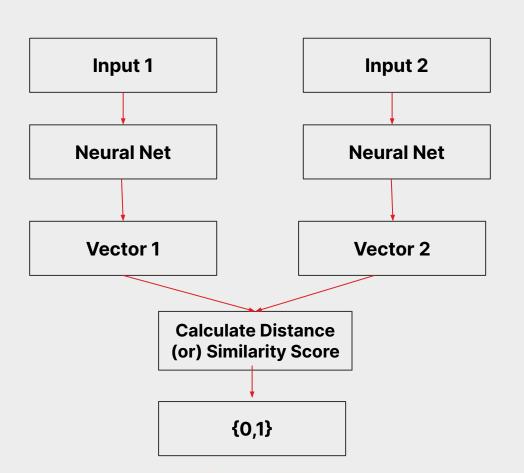
အဆင်ပြေကြပါစေ။

**Example Query** 

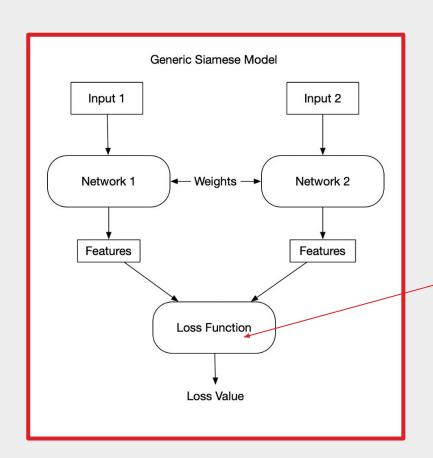




Same Weight,
Same Model Architecture







**Contrastive Loss Triplet Loss** 



### **Triplet Loss**

$$\sum_{i}^{N} \left[ \|f(x_{i}^{a}) - f(x_{i}^{p})\|_{2}^{2} - \|f(x_{i}^{a}) - f(x_{i}^{n})\|_{2}^{2} + \alpha \right]$$



#### **Contrastive Loss**

$$(0, \frac{1}{2})$$
  $(1-1)$ 

# **Creating Dataset (Contrastive Loss Setting)**

MultiLingual Semi Syllable Break for two Selected Sentence for each category

```
sentence11 = syllable_break("အား လုံး ပဲ ဘေး ရန် ကင်း ကွာ ကိုယ် စိတ် နှစ် ဖြာ ကျန်း မာ ချမ်း သာ ကြ ပါ စေ")
sentence12 = syllable_break("တ ဝါ တွင်း လုံး ကိုယ် ထက် ကြီး တဲ့ မ အ လ အ ဘ ကို မ အေ လုံး မ အေ လုံး နဲ့ ဆဲ ခဲ့ သ မျှ ဝါ ကျွက် မှာ ဝန် ချ တောင်း ပန် ပါ တယ် အ ဘ မ အ လ ဗွေ မ
print(sentence11)
print(sentence12)

□ အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ စ် ဖြာ ကျန်း မာ ချ မ်း သာ ကြ ပါ စေ
တ ဝါ တွင်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ ဘ ကို မ အေ လုံး မ အေ လုံး နဲ့ ဆဲ ခဲ့ သ မျှ ဝါ ကျွ က် မှာ ဝ န် ချ တော င်း ပ န် ပါ တ ယ် အ ဘ မ အ လ ဗွေ မ ယူ ပါ နဲ့
```

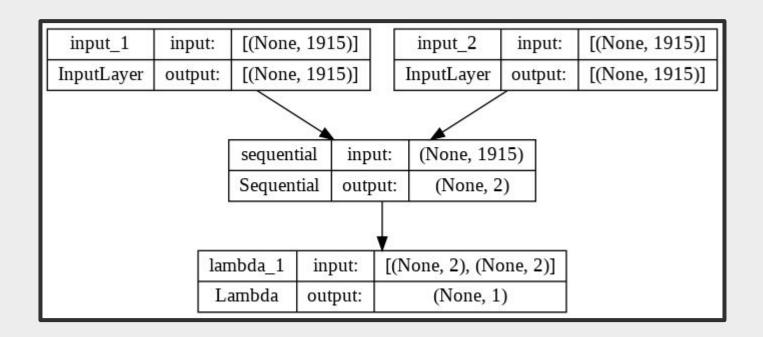


	sentence1	sentence2	labels
100	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	alp ဗို လ် ချူ ပ် တေ လ မ်း စ ဉ် သ ည် ရ ခို င်	1
101	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	1
102	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	တော် လှ န် ရေး ဟာ ဘ ယ် သူ က မှ စ တ င် ပြု ပြ င	1
103	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	ကို ဗ စ် ပိုး အ သ စ် ပေါ် လာ လို့ ဥ ရော ပ နဲ့	1
104	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	breaking news 🎓 တ န့် ဆ ည် ရဲ၂ ဦး သေ န တ်	1
*	No.	1000	2208
395	တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	လ မ်း ပေါ် မှာ ပ န်း လို က် ရော င်း တဲ့ က လေး	0
396	တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	ယ ခု ပို့ စ် သ ည် အ ထိ တ် တ လ န့် ဖြ စ် စေ	0
397	တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	ymba ဆို တာ အ ြ က မ်း ဖ က် အ ဖွဲ ့ အ စ ည်း ြ က	0
398	တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	ပြ ည် ထော င် စု မ ပြို ကွဲ ရေး တို င်း ရ င်း သ	0
399	တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	2 5 . 8 . 2 0 2 1 ငွေ ဖြ ည့် က ဒ် လ က် ကား ဈေ	0
300 rd	ows × 3 columns	\$1.00 (1.00 to 1.00 to	

#### Data Amount = nways\*(nways\*kshots)

		labels
အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	မြ န် မာ အ ရေး un ရဲ့ဆုံး ဖြ တ် ချ က် အ မှ န်	0
အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	ဒီ ဂ ရု ထဲ က အ ဘ မ င်း အော င် လှို င် ကို ချ စ	0
အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	ဪ ကွီး ဖြိုး ရ ယ် နှ စ် သ စ် ဆ န်း ရ င် မြ	0
အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	ဒေါ် စု ဆို တဲ့ ဒီ မို ဂ ရေ စီ ခေါ င်း ဆော င်	0
အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ	စ စ် တ ပ် ကော င်း ကြော င်း သိ ချ င် ရ င် မြ ဝ	0
		***
တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	လ မ်း ပေါ် မှာ ပ န်း လို က် ရော င်း တဲ့ က လေး	0
တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	ယ ခု ပို့ စ် သ ည် အ ထိ တ် တ လ န့် ဖြ စ် စေ	0
တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	ymba ဆို တာ အ ြ က မ်း ဖ က် အ ဖွဲ ့ အ စ ည်း ြ က	0
တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	ပြ ည် ထော င် စု မ ပြို ကွဲ ရေး တို င်း ရ င်း သ	0
တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	2 5 . 8 . 2 0 2 1 ငွေ ဖြ ည့် က ဒ် လ က် ကား ဈေ	0
	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ	အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ ဒီ ဂ ရု ထဲ က အ ဘ မ င်း အော င် လှို င် ကို ချ စ အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ သော် ကွီး ဖြိုး ရ ယ် နှ စ် သ စ် ဆ န်း ရ င် မြ အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ ဒေါ် စု ဆို တဲ့ ဒီ မို ဂ ရေ စီ ခေါ င်း ဆော င် အား လုံး ပဲ ဘေး ရ န် က င်း ကွာ ကို ယ် စိ တ် နှ စ စ် တ ပ် ကော င်း ကြော င်း သိ ချ င် ရ င် မြ ဝ တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ လ မ်း ပေါ် မှာ ပ န်း လို က် ရော င်း တဲ့ က လေး တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ ယ ခု ပို့ စ် သ ည် အ ထိ တ် တ လ န့် ဖြ စ် စေ တ ဝါ တွ င်း လုံး ကို ယ် ထ က် ကြီး တဲ့ မ အ လ အ ymba ဆို တာ အ ြ က မ်း ဖ က် အ ဖွဲ ့ အ စ ည်း ြ က ပြ ည် ထော င် စု မ ပြု ကွဲ ရေး တို င်း ရ င်း သ

## **Sample Model Architecture**



### **Testing**

```
======= 1 - 1s 776ms/step
1/1 [===========================] - 1s 767ms/step
ပွဲမြန်မြန်ပီးချင်ရင် ပလက်ဖောင်းမင်းသားသီချင်းနားထောင်ကြ သိပ်မတွေးနဲ့၊သိပ်မမေးနဲ့အချိန်ကုန်တယ် လေကုန်တယ် ငွေသာအရင် မ စ ပါ
                              - 1s 767ms/step
True 0
Predicted 0
    [======] - 1s 776ms/step
မြောင်က ဘာပြောလဲ မအလ ကို
True 0
Predicted 0
1/1 [======] - 1s 800ms/step
1/1 [======] - 1s 775ms/step
မအူဝဲ်ကိုပြေတီဦး ကိုလူမင်း မလွှတ်ပါ နောက်အမှုကပ်ထားတာရှိနေလိုပါ 😖
True 0
Predicted 0
    1/1 [======= - - 1s 763ms/step
ခန္ဓာကိုယ်ခုခံစနစ်ကအဆုတ်ကိုတိုက်ခိုက်တာသွေးကြဲဆေး ကြက်သွန်ဖြူ နနွင်း သံလွင်ဆီ ရှောက် သံပုရာ သုံးကြပါ
True 0
Predicted 0
```

f:x -> [0,1] i.e.

f(x)
= orgmox cos (o,x)

aeA

### References

- 1. Photos: Google
- 2. <a href="https://trishalaneeraj.github.io/2017-12-22/semantic-entailment">https://trishalaneeraj.github.io/2017-12-22/semantic-entailment</a>
- 3. <a href="https://kishaloyhalder.github.io/pdfs/tars\_coling2020.pdf">https://kishaloyhalder.github.io/pdfs/tars\_coling2020.pdf</a>
- 4. <a href="https://towardsdatascience.com/how-are-logistic-regression-ordinary-least-squares-regression-related-1deab32d79f5">https://towardsdatascience.com/how-are-logistic-regression-ordinary-least-squares-r
- 5. Activation Functions, <a href="https://sefiks.com/tag/activation-function/">https://sefiks.com/tag/activation-function/</a>
- 6. Yin et al., "Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach." arXiv preprint arXiv:1909.00161 (2019)



# ကျေးဇူး Thank You Gracias شکرا لك



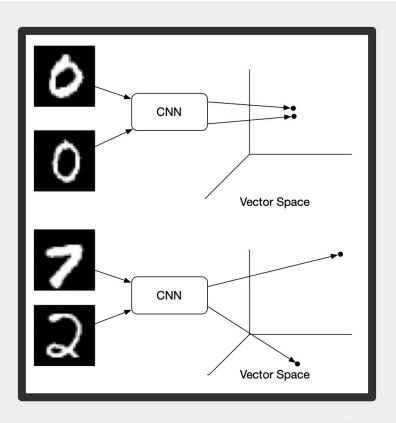




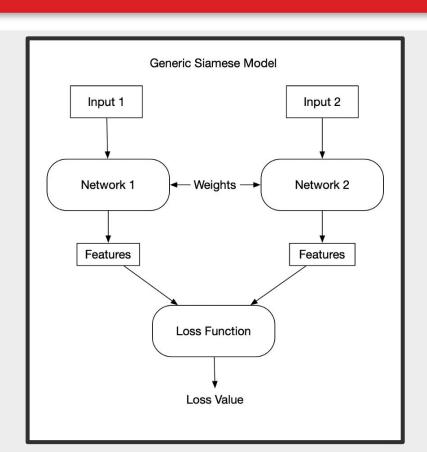
# **Extra Slides**



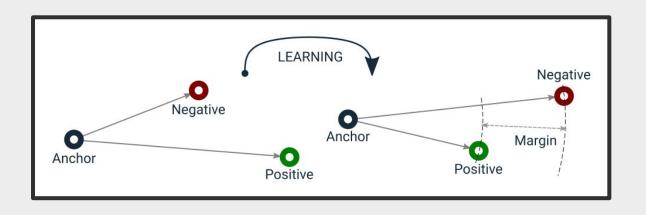
### Siamese Network











Ref: https://towardsdatascience.com/triplet-loss-advanced-intro-49a07b 7d8905



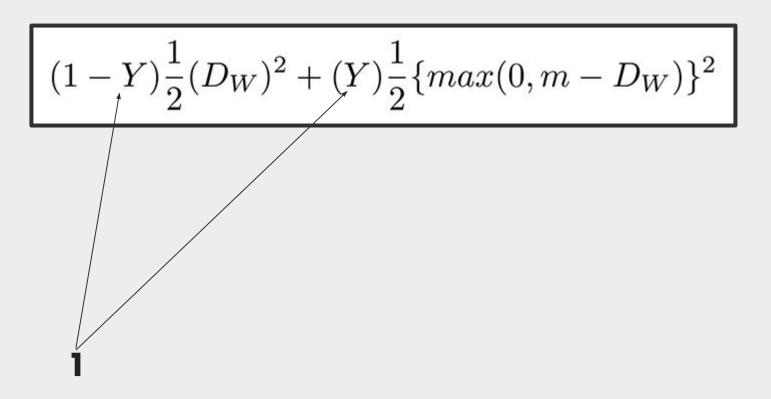
$$\sum_{i=1}^{N} \left[ \|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha \right]$$



$$(1-Y)\frac{1}{2}(D_W)^2 + (Y)\frac{1}{2}\{max(0, m-D_W)\}^2$$

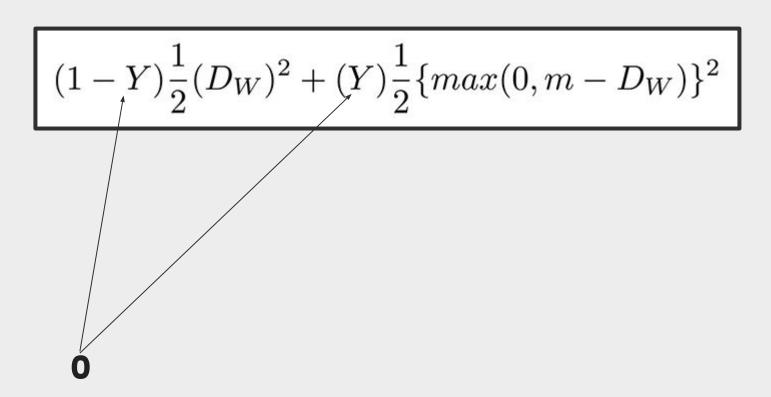


### When Y is 1





### When Y is 0





## **Anchor, Positive Pair**

Positive Anchor	Positive	Label

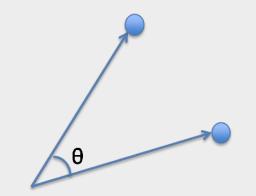


# **Anchor, Negative Pairs**

Negative Anchor	Negative	Label



$$sim(A, B) = cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$





1

#### **Positive Anchor**

### Input

$$\cos \theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$$
$$\|\vec{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2}$$
$$\|\vec{b}\| = \sqrt{b_1^2 + b_2^2 + b_3^2 + \dots + b_n^2}$$

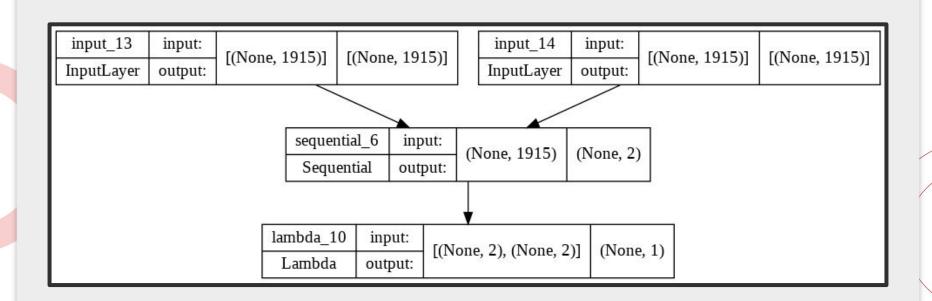
$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \sqrt{\sum_{i=1}^{n} B_i^2}}$$

$$sim(A, B) = cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|}$$



yi = argmax (sim(A, Input))



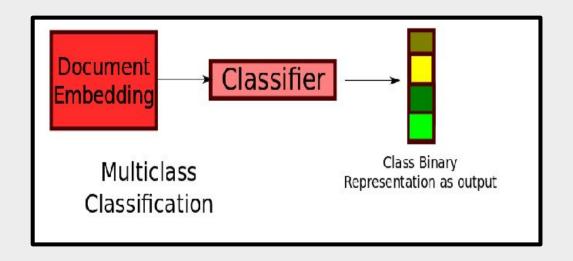


### **Harmonic Mean**

$$H = \frac{\frac{n}{1}}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \dots + \frac{1}{x_n}}$$



### Cont'd





Ref: Wai et al., Simotion, Emotion in Text