Classical Machine Learning coding:

1. KNN using Python/NumPy
2. Logistic Regression using Python/NumPy
3. Linear Regression using Python/NumPy
4. K-Mean using Python/NumPy
5. PCA using Python/NumPy
6. SVD using Python/NumPy

Deep Learning coding:

1. TF-IDF
2. 1- or 2-layer neural network using Python/NumPy
3. RNN and LSTM using Python/NumPy
4. LSA, LDA and PLSA topic modelling

NLP:

1. Embeddings
2. Classification models; LSTM, GRU, RNN, and Transformer encoder or decoder.
3. Topic Modelling
4. Siamese Networks; Contrastive loss, triplet loss, Quadruplet loss
5. A few Shot Short Learning (coding)
6. Attention Mechanism (self vs Simple attention)
7. Translation model: Seq-2-seq LSTM, Transformer model
8. LLM: BERT, XML Roberta, Electra, and GPT-2
9. Fine tuning and Transfer Learning (coding)
10. Summarization: 1. Abstractive 2. Extractive summarization (coding)
11. LLM+RHFL concept (coding)
12. Inference Optimization: XLA Complier, Pruning, Quantization, TensorRT and Distillation (coding)

Named entity, intent entity, sentiment anslysis, spell checker, next word predictions --- BERT

Proper defination:

Embedding : vector representations are learned during training and can capture semantic relationship between words.

It is first learnable latyers of neural network, it takes tensor of integer indices and return dense tensor of dense layer.

# vocabulary size = 1000 and **embedding dimension = 200**

embedding = nn.Embedding(1000, 200)

# Input tensor of shape (**batch\_size = 2, sequence\_length =3**)

inputs = torch.LongTensor([[1, 2, 3], [4, 5, 6]])

# Pass the inputs through the embedding layer

embedded\_inputs = embedding(inputs)

# output tensor shape = (**batch\_size, sequence\_length, embedding\_dim**)

print(embedded\_inputs.shape)

torch.Size([**2**, **3**, **200**])

embedded\_inputs[0].shape

torch.Size([3, 200])

===============================

Embedding initialization :

orthogonal(svd) matrix \* matrix

* Trnspose will give diagonal as 1 and remaining zero.
* preserve the gradients over many timesteps

Uniform can not go outside a range like -1 to 1

* random can give any random value in weight

Embedding regularizer : L1 and L2

Weight?

Word2vec – genism example?

Glove

Fastag -- Multi lingua? example

Benfits?

Why embedding layers?

Rather then using Bert embedding we use the bert tokenizer

Example?

--------------------------

How transformer learn different at different layers?