Use Cases

Classification

Chatbots

Summarization

Evolution of NLP

Tf-Idf, Word count- Naïve Bayes, RF

Word embeddings – Glove, Fasttext, Word2Vec

Bert, ULMFiT, etc.

About Speaker

CTO @ SmartNomad

AI strategy – Advisor/Mentor to enterprises and startups.

Author, Deep Learning

Speaker

Top 1% in Kaggle competitions (Competition Expert)

How can NLP add value to an organization

Background	Solution Design
A B2C Fashion e-commerce company aiming to improve customer complaint resolution process	NLP Based algorithm can extract attributes from raw customer complaint text and highlight department(s) needed to resolve the problem. A near real-time notification is sent to relevant teams to implement corrective action.
A B2C travel company trying to increase customer base by serving real time recommendations and providing online support	A Chatbot based service that derives context from a user's input to an UI and telemetry data. This helps provide personalized and experiences thereby driving significant impact on customer acquisition, engagement and retention
A large enterprise attempting to improve internal org operations by automating policy related questions (onboarding for new joiners, healthcare, IT etc.)	An NLP based solution that generates easy-to-manage and exhaustive checklists by interpreting the syntax and semantic analysis of questions.

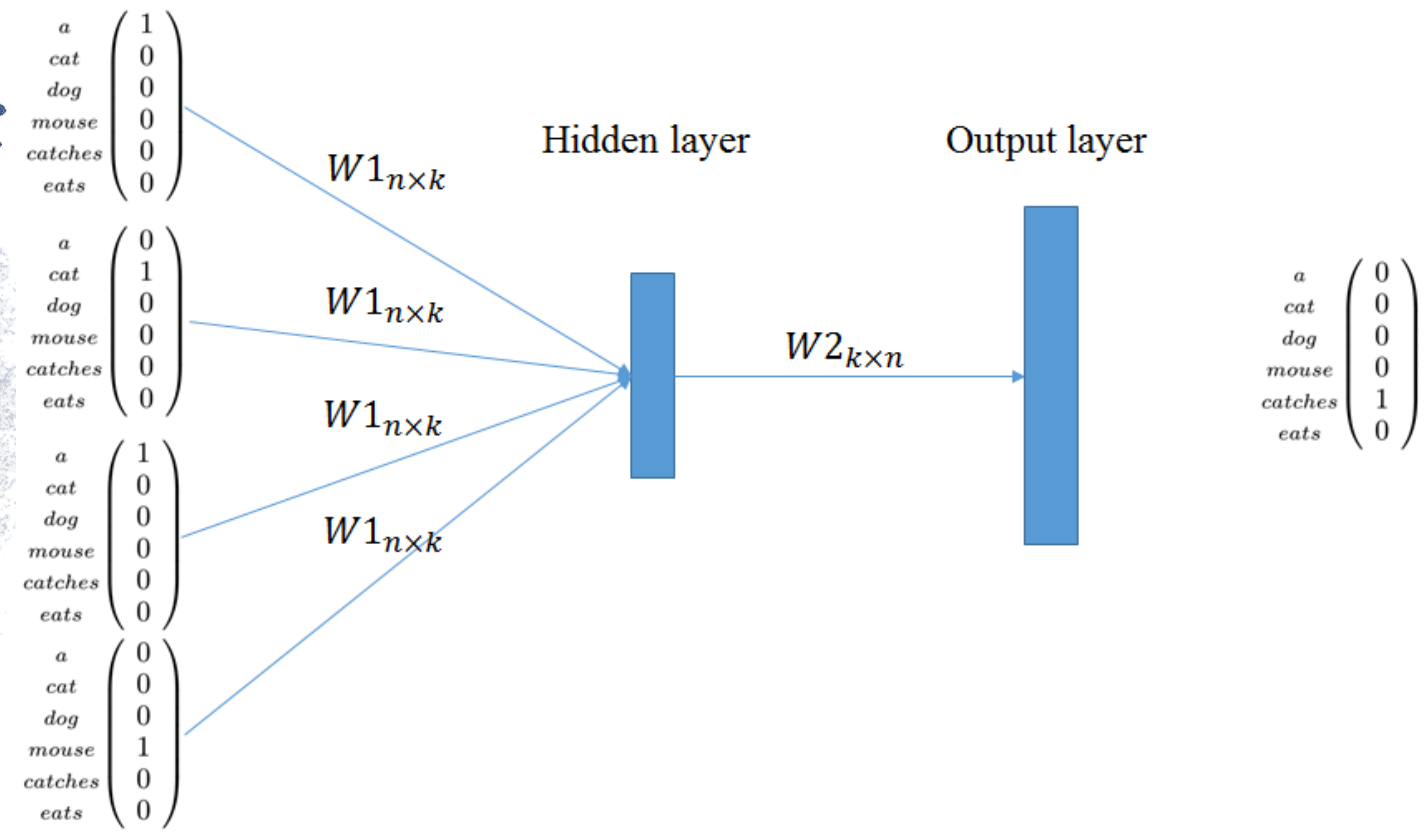
Evolution of NLP

BOW, TF-IDF, hand crafted features - Naive bayes, Random Forest, DL (2001-2012)

Word Embeddings like Word2Vec, Glove, Fasttext – Deep learning algotihms (2013-2018)

Transfer learning based approaches Bert, ULMFiT, ELMo. (2018 *)

BOW, TF-IDF


a cat dog mouse catches eats
1 0 0 0 0 0
0 1 0 0 0 0
1 0 0 0 0 0
0 0 0 1 0 0
$W1_{n\times k}$
$W1_{n\times k}$
$W1_{n\times k}$
$W1_{n\times k}$
Hidden layer
$W2_{k\times n}$
Output layer
0 0 0 0 1 0

Word embeddings

A 4-dimensional embedding

cat =>	1.2	-0.1	4.3	3.2
mat =>	0.4	2.5	-0.9	0.5
on =>	2.1	0.3	0.1	0.4
...	...			

Glove
FastText
Word2vec

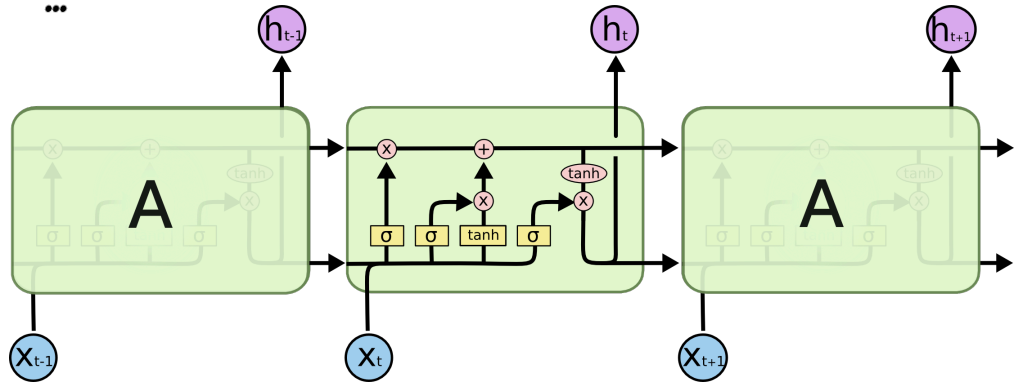


Figure 5: An LSTM network (Source: Chris Olah)

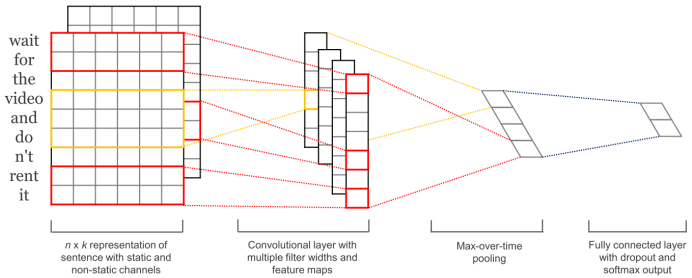
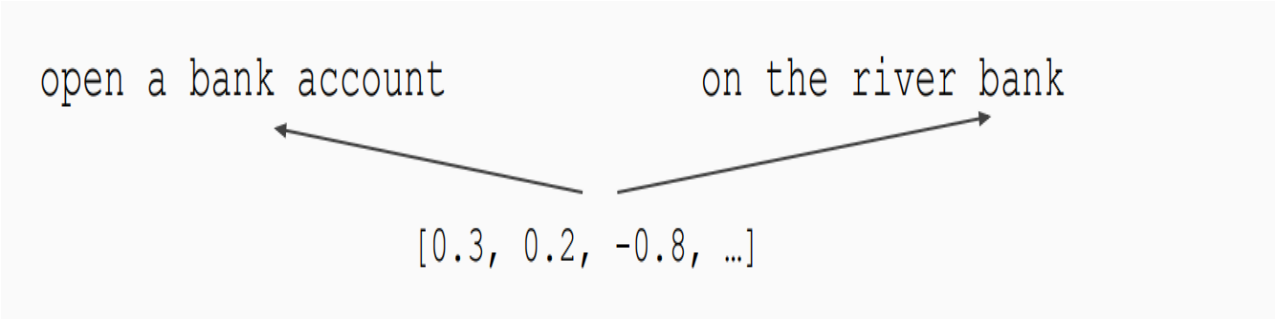


Figure 6: A convolutional neural network for text (Kim, 2014)

Missing Context

Words have different contexts based on the underlying context



Transfer Learning

1. Unsupervised Learning – LM

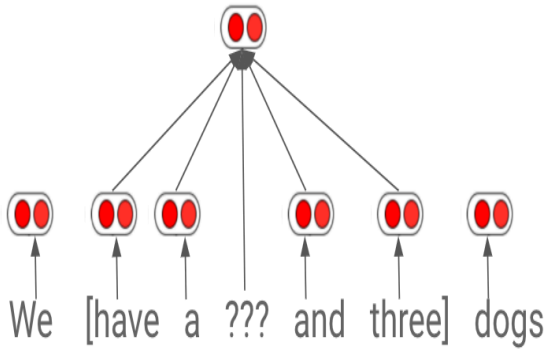
2. Fine tuning – ULMFiT, Bert

Transfer Learning

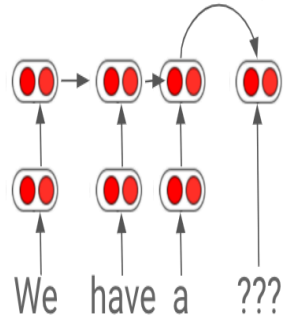
1. What is Language Model
2. Advantages of Language Model
3. Different approaches to train a LM

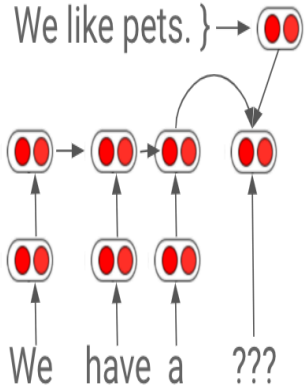
Language model

word2vec, Mikolov et al (2013)



ELMo, Peters et al. 2018, ULMFiT (Howard & Ruder 2018), GPT (Radford et al. 2018)





Skip-Thought (Kiros et al., 2015)

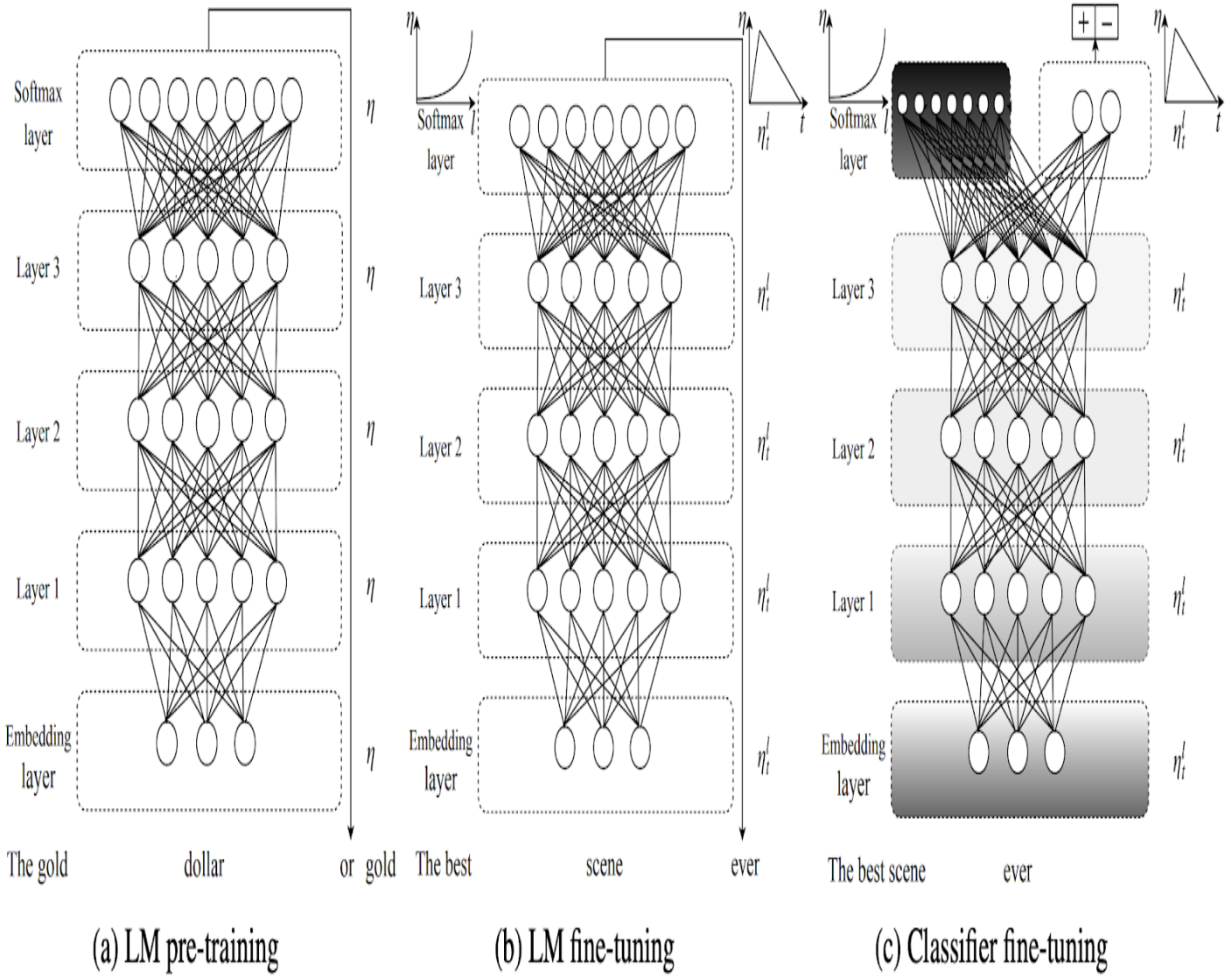
BERT, Devlin et al 2019

Finetuning LM

Two algorithms which made significant progress.

- ULMFiT
- Bert

UMLFit



(Howard and Ruder, ACL 2018)

ULMFiT

Key steps to train your own ULMFiT model

1. Train the LM model on a large corpus of data like Wikipedia. For major languages you should be able to download pre trained weights.

2. Finetune the LM model for the specific domains like Medical documents, movie scripts

3. Build a classifier, QA kind of applications on top of that

Bert-LM

Key contribution to train LM

Sentence A = The man went to the store.
Sentence B = He bought a gallon of milk.
Label = IsNextSentence

Sentence A = The man went to the store.
Sentence B = Penguins are flightless.
Label = NotNextSentence

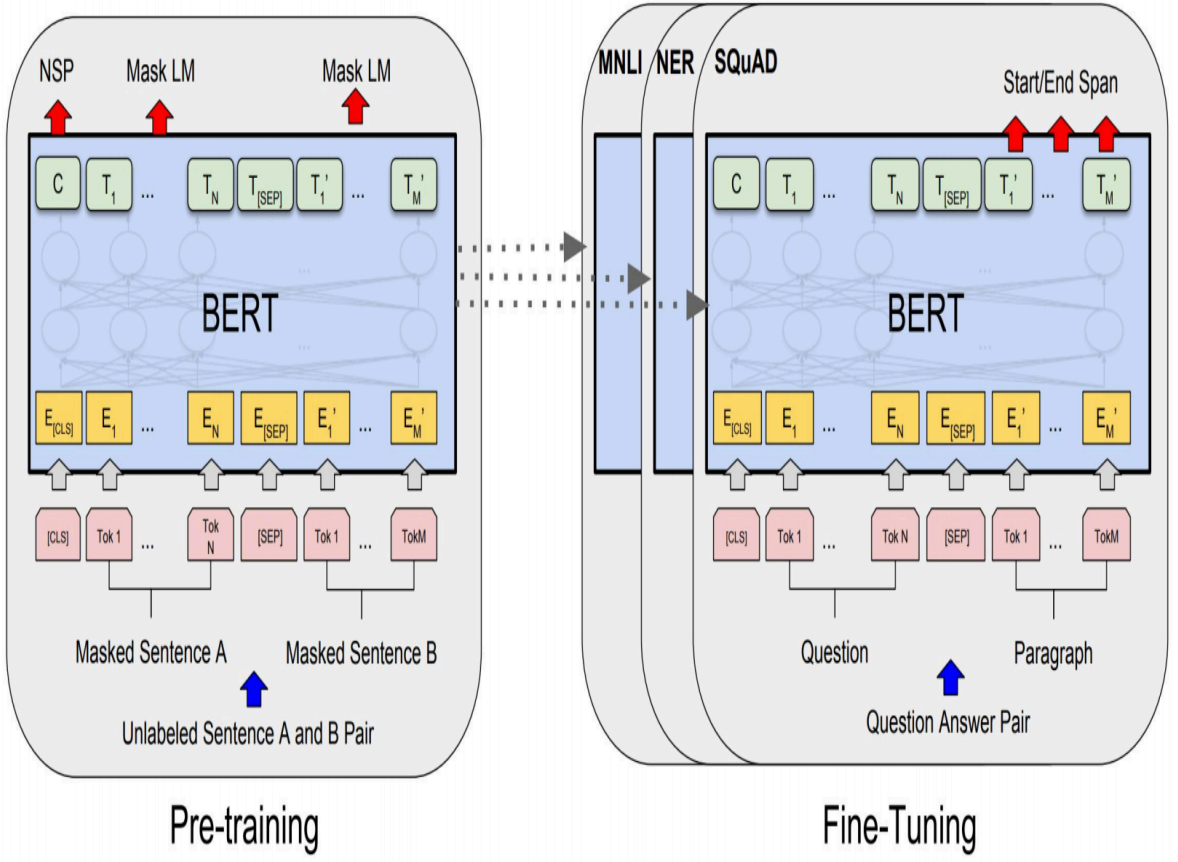
Bert-LM

Key contribution to train LM

Sentence A = The man went to the store.
Sentence B = He bought a gallon of milk.
Label = IsNextSentence

Sentence A = The man went to the store.
Sentence B = Penguins are flightless.
Label = NotNextSentence

Bert-Finetuning



Pros and Cons of Bert

Pros

- Achieved state of art results in different NLP tasks like classification, summarization, Q&A.
- Easy to fine tune for custom use cases.
- Since the architecture relies on transformers, it can be easily parallelized unlike RNN- variants.

Cons

- Training a BERT model from scratch could take days to weeks depending on the kind of GPU's used.
- Efficient finetuning approaches are still being researched as its still a new technique.
- Inference on production systems could be challenging due to the size of the model.

Thank you

CONTACT DETAILS –
9003271410

LINKEDIN