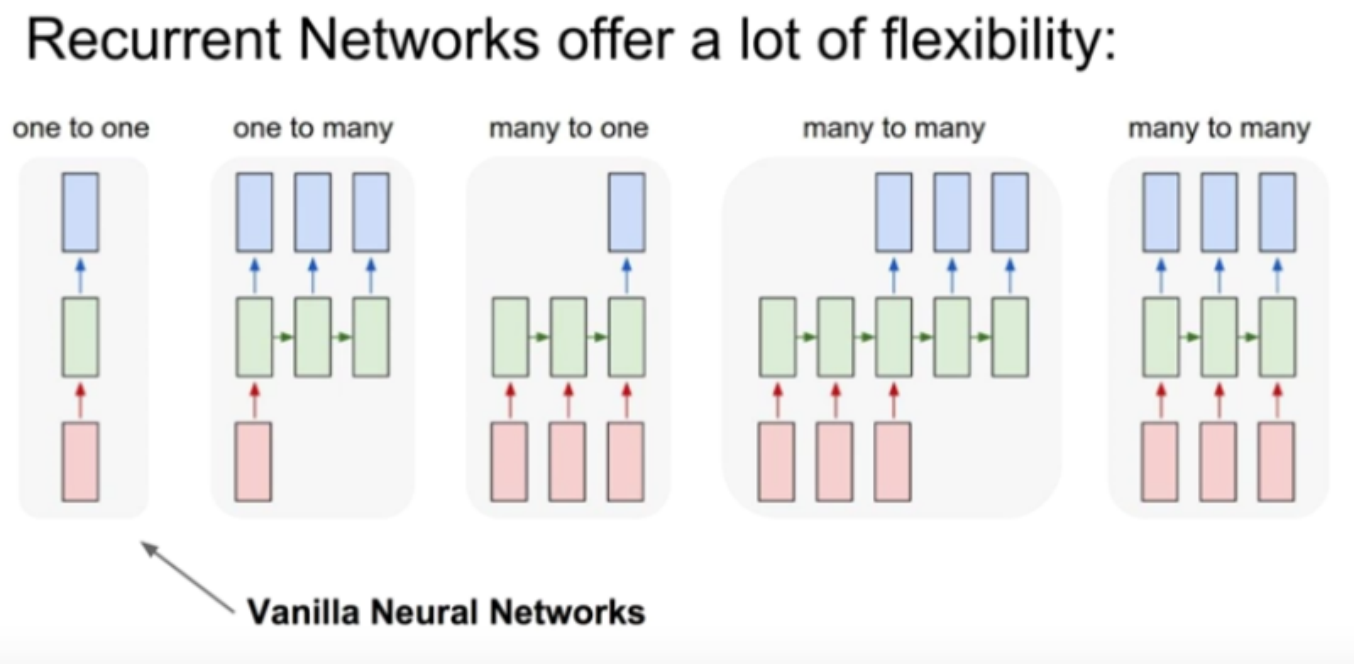
Today's topic is RNN

RNN can handle diverse task

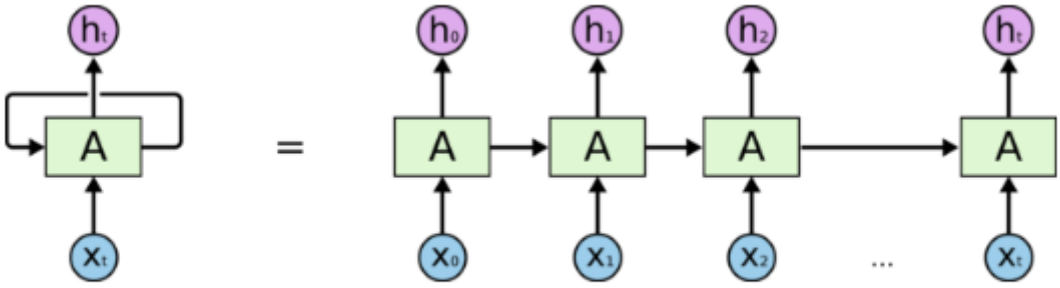
Today, we implement below models

1. many-to-many RNN
2. many-to-one RNN



1. Word-RNN(Recurrent Neural Network)

Many-to-Many RNN: we are implemnt the text generation rnn



An unrolled recurrent neural network.

In [1]:

```
# import pakages
import os
!pip install torchtext==0.8.1
import torchtext as torchtext
import torch
```

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheel
s/public/simple/
Collecting torchtext==0.8.1
  Downloading torchtext-0.8.1-cp37-cp37m-manylinux1_x86_64.whl (7.0 MB)
     |████████████████████████████████| 7.0 MB 19.9 MB/s
Requirement already satisfied: requests in /usr/local/lib/python3.7/dist-packages
(from torchtext==0.8.1) (2.23.0)
Collecting torch==1.7.1
  Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
     |████████████████████████████████| 776.8 MB 17 kB/s
Requirement already satisfied: tqdm in /usr/local/lib/python3.7/dist-packages (fro
m torchtext==0.8.1) (4.64.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages (fr
om torchtext==0.8.1) (1.21.6)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.7/dist-
packages (from torch==1.7.1->torchtext==0.8.1) (4.1.1)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in /usr/loc
al/lib/python3.7/dist-packages (from requests->torchtext==0.8.1) (1.24.3)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist
-packages (from requests->torchtext==0.8.1) (2022.9.24)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packa
ges (from requests->torchtext==0.8.1) (2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-
packages (from requests->torchtext==0.8.1) (3.0.4)
Installing collected packages: torch, torchtext
  Attempting uninstall: torch
    Found existing installation: torch 1.12.1+cu113
    Uninstalling torch-1.12.1+cu113:
      Successfully uninstalled torch-1.12.1+cu113
  Attempting uninstall: torchtext
    Found existing installation: torchtext 0.13.1
    Uninstalling torchtext-0.13.1:
      Successfully uninstalled torchtext-0.13.1
ERROR: pip's dependency resolver does not currently take into account all the pack
ages that are installed. This behaviour is the source of the following dependency
 conflicts.
torchvision 0.13.1+cu113 requires torch==1.12.1, but you have torch 1.7.1 which is
incompatible.
torchaudio 0.12.1+cu113 requires torch==1.12.1, but you have torch 1.7.1 which is
 incompatible.
Successfully installed torch-1.7.1 torchtext-0.8.1
```

In [2]:

```
import warnings
warnings.filterwarnings(action='ignore')
```

1-1. Data preprocessing

In [3]:

```
#Simple sentence data.
sentence = "All grownup were once children, although few of them remember it".split()
#Make the vokabulary
vocab = list(set(sentence))
print(vocab)
```

```
['children,', 'few', 'although', 'of', 'grownup', 'All', 'remember', 'it', 'once',
'were', 'them']
```

In [4]:

```
#We use the word indexing,it is converted to one-hot encoding inside the model.
#This dict transfroms word to index(number)
word2index = {tkn: i for i, tkn in enumerate(vocab, 1)}
word2index['<unk>']=0
print(word2index)
```

```
{'children,': 1, 'few': 2, 'although': 3, 'of': 4, 'grownup': 5, 'All': 6, 'rememb
er': 7, 'it': 8, 'once': 9, 'were': 10, 'them': 11, '<unk>': 0}
```

In [5]:

```
#This dict transfroms index to number(number)
index2word = {v: k for k, v in word2index.items()}
print(index2word)
```

```
{1: 'children,', 2: 'few', 3: 'although', 4: 'of', 5: 'grownup', 6: 'All', 7: 'rem
ember', 8: 'it', 9: 'once', 10: 'were', 11: 'them', 0: '<unk>'}
```

In [6]:

```
#Function to make the input data(X) and the labels(Y)
def build_data(sentence, word2index):
  encoded = [word2index[token] for token in sentence] # transforms word to index
  input_seq, label_seq = encoded[:-1], encoded[1:] # Split the input sequence and label sequence
  input_seq = torch.LongTensor(input_seq).unsqueeze(0)
  label_seq = torch.LongTensor(label_seq).unsqueeze(0)
  return input_seq, label_seq, encoded
```

In [7]:

```
x, y, encoded = build_data(sentence, word2index)
```

In [8]:

```
print(f'Input data: {x} sentence: All grownup were once children, although few of them remember
Wnlabes: {y} sentence: grownup were once children, although few of them remember it')
```

```
Input data: tensor([[ 6,  5, 10,  9,  1,  3,  2,  4, 11,  7]]) sentence: All grown
up were once children, although few of them remember
labes: tensor([[ 5, 10,  9,  1,  3,  2,  4, 11,  7,  8]]) sentence: grownup were o
nce children, although few of them remember it
```

Word Embedding

In order to handle text data, a process of changing word to a number through embedding is required.

There are so many embedding methods, and one hot encoding that we learned is also one of the embedding methods.

Label Encoding

Food Name	Categorical #	Calories
Apple	1	95
Chicken	2	231
Broccoli	3	50

→

One Hot Encoding

Apple	Chicken	Broccoli	Calories
1	0	0	95
0	1	0	231
0	0	1	50

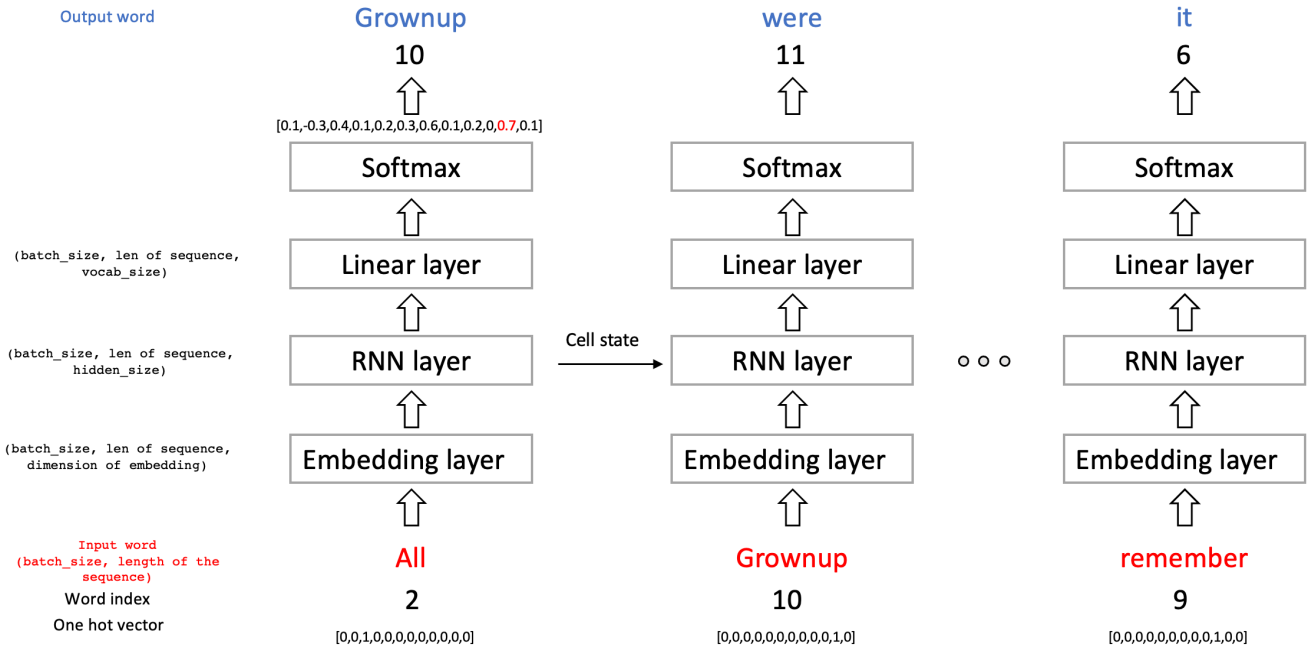
In [9]:

```
embedding_function = torch.nn.Embedding(num_embeddings=len(word2index), embedding_dim = 5)
embedding_function(x)
```

Out[9]:

```
tensor([[[ 0.3908,  0.1375, -0.0215,  0.4859, -1.1237],
         [-0.6476, -0.5283, -1.0195, -0.3965,  0.3818],
         [-1.0504, -0.2549,  0.5693,  0.3564, -1.3894],
         [-1.5161, -0.5830,  0.2423, -0.3684, -0.0307],
         [ 1.1377,  1.8763,  0.2513, -1.0658, -0.1325],
         [-0.0223, -0.5401, -0.4157,  0.4079, -0.5631],
         [ 0.3334, -0.1141, -0.3307,  0.8333,  0.7801],
         [ 1.1563,  0.5188, -0.2368,  1.4955,  1.5969],
         [ 0.6766, -1.3343,  0.8149,  0.1451, -0.3924],
         [-0.2526, -1.7122, -0.4085, -1.0105, -0.6752]]],
       grad_fn=<EmbeddingBackward>)
```

1-2. Implenment the Many-to-Many RNN model



In [10]:

```
#https://pytorch.org/docs/stable/generated/torch.nn.RNN.html
class RNN_model(torch.nn.Module):
  #vocab_size = size of the using word
  #hidden_size = size of the RNN's output
  def __init__(self, vocab_size, input_size, hidden_size, batch_first=True):
    super(RNN_model, self).__init__()

    #Embedding layer
    self.embedding_layer = torch.nn.Embedding(num_embeddings=vocab_size,embedding_dim=input_size
)

    #RNN layer
    self.rnn_layer = torch.nn.RNN(input_size=input_size, hidden_size=hidden_size, batch_first=ba
tch_first)

    #linear layer
    self.linear = torch.nn.Linear(hidden_size, vocab_size)

  def forward(self, x):
    #1. Embedding layer
    #size of the data: (batch_size, length of the sequence)
    #-> (batch_size, lenth of the sequence, dimension of embedding)
    y = self.embedding_layer(x)

    #2. RNN layer
    #size of the data: (batch_size, length of the sequence, dimension of embedding)
    #-> y: (batch_size, length of the sequence, hidden_size), hidden: (1, batch_size, hidden_siz
e)
    y, hidden = self.rnn_layer(y)

    #3. Linear layer
    #size of the data: (batch_size, length of the sequence, hidden_size)
    #-> (batch_size, length of the sequence, vocab_size)
    y = self.linear(y)

    #Size of the return value: (batch_size*lenth of the sequence, vocab_size)
    return y.view(-1, y.size(2))
```

1-3. Train the RNN model

In [11]:

```
#hyper parameter
vocab_size = len(word2index)
input_size = 5
hidden_size = 20
epochs = 100
```

In [12]:

```
model = RNN_model(vocab_size, input_size, hidden_size, batch_first=True)
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(params=model.parameters())
```

In [13]:

```
#This is the function to decode the model result into word.
decode = lambda y: [index2word.get(x) for x in y]
```

In [14]:

```
#Befor the learning, observe the result
output = model(x)
pred = output.softmax(-1).argmax(-1).tolist()

print(f'Law x: All grownup were once children, although few of them remember')
print(f'Result of the rnn model: {" ".join(decode(pred))}')
```

```
Law x: All grownup were once children, although few of them remember
Result of the rnn model: although remember remember of of <unk> remember children,
remember although
```

In [15]:

```
#Training
for epoch in range(1, epochs+1):
    optimizer.zero_grad()
    output = model(x)
    loss = loss_function(output, y.view(-1))
    loss.backward()
    optimizer.step()
    # Observe the result
    if epoch % 20 == 0:
        print(f"Epoch: {epoch}, Loss: {loss}")
        pred = output.softmax(-1).argmax(-1).tolist()
        print(" ".join(['All']+ decode(pred)))
        print()
```

```
Epoch: 20, Loss: 2.289281129837036
All although children, remember of of although of were remember although

Epoch: 40, Loss: 2.04744029045105
All children, were remember children, of few of were remember it

Epoch: 60, Loss: 1.7511638402938843
All it were remember children, of few of were remember it

Epoch: 80, Loss: 1.4044287204742432
All grownup were once children, of few of them remember it

Epoch: 100, Loss: 1.0769438743591309
All grownup were once children, although few of them remember it
```

2. Sentence classification RNN

Many to one RNN

In [16]:

```
USE_CUDA = torch.cuda.is_available()
DEVICE = torch.device("cuda" if USE_CUDA else "cpu")
print("Using device:", DEVICE)
```

```
Using device: cuda
```

2-1. Load the IMDB data

IMDB data consists of text data about movie reviews.

Labels are whether reviews are positive or negative.

IMDB is data for learning a classification model that classifies movie reviews as positive or negative.

In [17]:

```
#It is an instructor for transforming text data into a tensor.
# https://pytorch.org/text/_modules/torchtext/data/field.html
TEXT = torchtext.data.Field(sequential=True, batch_first=True, lower=True)
LABEL = torchtext.data.Field(sequential=False, batch_first=True)
```

In [18]:

```
trainset, testset = torchtext.datasets.IMDB.splits(TEXT, LABEL)
```

```
downloading aclImdb_v1.tar.gz

aclImdb_v1.tar.gz: 100%|██████████| 84.1M/84.1M [00:09<00:00, 8.42MB/s]
```

In [19]:

```
print(f'Length of trainset: {len(trainset)} Length of testset: {len(testset)}')
```

```
Length of trainset: 25000 Length of testset: 25000
```

In [20]:

```
#Observe the data
pos_data = vars(trainset[0])
neg_data = vars(trainset[20000])
print(f"--Positive data-- ₩n Sentence: {' '.join(pos_data['text'])}₩n Label: {pos_data['label']
}")
print()
print(f"--Negative data-- ₩n Sentence: {' '.join(neg_data['text'])}₩n Label: {neg_data['label']
}")
```

```
--Positive data--
 Sentence: such a film of beauty that it's hard to describe. maybe it's the absenc
e of superfluous dialogue, or maybe it's the absolutely stellar soundtrack, or may
be it's just meena mumari's feet, but it's a joy to watch this movie again and aga
in. i've never seen another indian movie that comes close to it, and few from any
country rival its perfection.
 Label: pos

--Negative data--
 Sentence: have to admit, this version disgraces shakespeare upfront! none can act
except the nurse who was my fav! juliet had good skills as a teen but she can't gi
ve emotional depth to her lines and we really can never connect to her. she's wors
e doing the scene when she is contemplating drinking the sleeping potion...god sto
p whining! i would have poured it in her mouth to shut her up! anthony andrews...y
ikes! considering his other great movies (brideshead revisited, ivanhoe, scarlet p
impernel), he's quite a shocker in this one. and don't get me started on romeo...p
uhleasssssee! it's still good to see if you're on the hunt to see every romeo and
juliet ever made in the history of film. olivia and leonard's version is still the
best, followed by leslie howard's version and then the current leo and clare!
 Label: neg
```

In [21]:

```
#Make the vocabulary set
TEXT.build_vocab(trainset, min_freq=5)
LABEL.build_vocab(trainset)
```

In [22]:

```
vocab_size = len(TEXT.vocab)
print(f'Size of the vocabulary set: {vocab_size}')
```

```
Size of the vocabulary set: 46159
```

In [23]:

```
#Observe the words
TEXT.vocab.stoi
```

Out[23]:

```
defaultdict(<bound method Vocab._default_unk_index of <torchtext.vocab.Vocab objec
t at 0x7f3d67befb10>>,
            {'<unk>': 0,
             '<pad>': 1,
             'the': 2,
             'a': 3,
             'and': 4,
             'of': 5,
             'to': 6,
             'is': 7,
             'in': 8,
             'i': 9,
             'this': 10,
             'that': 11,
             'it': 12,
             '/><br': 13,
             'was': 14,
             'as': 15,
             'for': 16,
             'with': 17,
             'but': 18,
             'on': 19,
             'movie': 20,
             'his': 21,
             'are': 22,
             'not': 23,
             'film': 24,
             'you': 25,
             'have': 26,
             'he': 27,
             'be': 28,
             'at': 29,
             'one': 30,
             'by': 31,
             'an': 32,
             'they': 33,
             'from': 34,
             'all': 35,
             'who': 36,
             'like': 37,
             'so': 38,
             'just': 39,
             'or': 40,
             'has': 41,
             'her': 42,
             'about': 43,
             "it's": 44,
             'some': 45,
             'if': 46,
             'out': 47,
             'what': 48,
             'very': 49,
             'when': 50,
             'more': 51,
             'there': 52,
             'she': 53,
             'would': 54,
             'even': 55,
             'good': 56,
```

```
'my': 57,
'only': 58,
'their': 59,
'no': 60,
'really': 61,
'had': 62,
'which': 63,
'can': 64,
'up': 65,
'were': 66,
'see': 67,
'than': 68,
'we': 69,
'-': 70,
'been': 71,
'into': 72,
'get': 73,
'will': 74,
'story': 75,
'much': 76,
'because': 77,
'most': 78,
'how': 79,
'other': 80,
'also': 81,
'first': 82,
'its': 83,
'time': 84,
'do': 85,
"don't": 86,
'me': 87,
'great': 88,
'people': 89,
'could': 90,
'make': 91,
'any': 92,
'/>the': 93,
'after': 94,
'made': 95,
'then': 96,
'bad': 97,
'think': 98,
'being': 99,
'many': 100,
'him': 101,
'never': 102,
'two': 103,
'too': 104,
'little': 105,
'where': 106,
'well': 107,
'<br': 108,
'way': 109,
'watch': 110,
'your': 111,
'it.': 112,
'did': 113,
'does': 114,
'them': 115,
'best': 116,
'movie.': 117,
```

```
 'know': 118,
 'seen': 119,
 'love': 120,
 'characters': 121,
 'character': 122,
 'movies': 123,
 'these': 124,
 'ever': 125,
 'still': 126,
 'over': 127,
 'should': 128,
 'films': 129,
 'such': 130,
 'plot': 131,
 'acting': 132,
 'while': 133,
 'show': 134,
 'go': 135,
 'those': 136,
 'off': 137,
 'better': 138,
 'film.': 139,
 'through': 140,
 "doesn't": 141,
 'say': 142,
 'something': 143,
 'why': 144,
 "i'm": 145,
 'makes': 146,
 "didn't": 147,
 'watching': 148,
 'back': 149,
 'scene': 150,
 'film,': 151,
 'real': 152,
 'find': 153,
 'new': 154,
 'movie,': 155,
 'few': 156,
 'actually': 157,
 'every': 158,
 'scenes': 159,
 'man': 160,
 'life': 161,
 'going': 162,
 'same': 163,
 'nothing': 164,
 '/>i': 165,
 'look': 166,
 'another': 167,
 'quite': 168,
 'lot': 169,
 'old': 170,
 'want': 171,
 'end': 172,
 'pretty': 173,
 'thing': 174,
 'seems': 175,
 'got': 176,
 '&': 177,
 "can't": 178,
```

```
 'before': 179,
 'take': 180,
 'years': 181,
 'part': 182,
 'actors': 183,
 'give': 184,
 'young': 185,
 'may': 186,
 'between': 187,
 'us': 188,
 "that's": 189,
 "i've": 190,
 'without': 191,
 'though': 192,
 'both': 193,
 'things': 194,
 'gets': 195,
 'big': 196,
 'around': 197,
 'here': 198,
 'thought': 199,
 'saw': 200,
 'director': 201,
 'almost': 202,
 'it,': 203,
 'now': 204,
 "isn't": 205,
 'always': 206,
 'must': 207,
 'come': 208,
 'own': 209,
 'work': 210,
 'whole': 211,
 'cast': 212,
 'horror': 213,
 'down': 214,
 'might': 215,
 "there's": 216,
 "he's": 217,
 'bit': 218,
 'least': 219,
 'probably': 220,
 'enough': 221,
 '"the': 222,
 'feel': 223,
 'last': 224,
 'original': 225,
 'am': 226,
 'since': 227,
 'rather': 228,
 'long': 229,
 'far': 230,
 'fact': 231,
 'kind': 232,
 'each': 233,
 'world': 234,
 'funny': 235,
 'found': 236,
 'anything': 237,
 'worst': 238,
 'comes': 239,
```

```
'having': 240,
'our': 241,
'making': 242,
'trying': 243,
'action': 244,
'right': 245,
'interesting': 246,
'done': 247,
'however,': 248,
'point': 249,
'believe': 250,
'looks': 251,
'guy': 252,
'put': 253,
'goes': 254,
'/>this': 255,
'family': 256,
'played': 257,
'main': 258,
'anyone': 259,
'series': 260,
'hard': 261,
"wasn't": 262,
'role': 263,
'especially': 264,
'music': 265,
'yet': 266,
'worth': 267,
'seem': 268,
'performance': 269,
'plays': 270,
'takes': 271,
'script': 272,
'watched': 273,
'sure': 274,
'looking': 275,
'during': 276,
'someone': 277,
'minutes': 278,
'different': 279,
'tv': 280,
'although': 281,
'woman': 282,
'set': 283,
'three': 284,
'away': 285,
'times': 286,
'shows': 287,
'maybe': 288,
'comedy': 289,
'girl': 290,
'left': 291,
'everything': 292,
'time.': 293,
'john': 294,
'once': 295,
'seeing': 296,
'simply': 297,
"you're": 298,
'american': 299,
'fun': 300,
```

```
'special': 301,
'completely': 302,
'everyone': 303,
'play': 304,
'true': 305,
'again': 306,
'reason': 307,
'read': 308,
'used': 309,
'well,': 310,
'need': 311,
'given': 312,
'until': 313,
'nice': 314,
'beautiful': 315,
'use': 316,
'high': 317,
'sense': 318,
'truly': 319,
'place': 320,
'idea': 321,
'--': 322,
'help': 323,
'version': 324,
'try': 325,
'less': 326,
'rest': 327,
'black': 328,
'money': 329,
'came': 330,
'job': 331,
'second': 332,
'dvd': 333,
'excellent': 334,
'recommend': 335,
'ending': 336,
'tell': 337,
'getting': 338,
'shot': 339,
'keep': 340,
'instead': 341,
'actor': 342,
'(and': 343,
'gives': 344,
'full': 345,
'said': 346,
'let': 347,
'half': 348,
'enjoy': 349,
'poor': 350,
'couple': 351,
'day': 352,
'himself': 353,
'playing': 354,
'definitely': 355,
'supposed': 356,
'audience': 357,
'early': 358,
'become': 359,
'along': 360,
'felt': 361,
```

```
'effects': 362,
'understand': 363,
'all,': 364,
'remember': 365,
'small': 366,
"couldn't": 367,
'book': 368,
'entire': 369,
'later': 370,
'absolutely': 371,
'liked': 372,
'together': 373,
'star': 374,
'against': 375,
'fan': 376,
'went': 377,
'wife': 378,
'next': 379,
'doing': 380,
'start': 381,
'perhaps': 382,
'year': 383,
'often': 384,
'hollywood': 385,
'certainly': 386,
'screen': 387,
'several': 388,
'(the': 389,
'time,': 390,
'2': 391,
"she's": 392,
'sort': 393,
'night': 394,
'waste': 395,
'human': 396,
'is,': 397,
'becomes': 398,
'short': 399,
'wonderful': 400,
'10': 401,
'seemed': 402,
'loved': 403,
'men': 404,
'kids': 405,
'₩x96': 406,
'piece': 407,
'war': 408,
'classic': 409,
'father': 410,
'.': 411,
'production': 412,
'house': 413,
'home': 414,
'camera': 415,
'else': 416,
'wanted': 417,
'hope': 418,
'women': 419,
'that,': 420,
'totally': 421,
'live': 422,
```

```
 'course': 423,
 'performances': 424,
 'lost': 425,
 'top': 426,
 'based': 427,
 "i'd": 428,
 'them.': 429,
 'tries': 430,
 'line': 431,
 'mind': 432,
 'video': 433,
 "you'll": 434,
 'final': 435,
 'able': 436,
 'called': 437,
 'wants': 438,
 'perfect': 439,
 'friends': 440,
 'gave': 441,
 "they're": 442,
 'under': 443,
 'death': 444,
 'already': 445,
 'despite': 446,
 'enjoyed': 447,
 'story,': 448,
 'finally': 449,
 'person': 450,
 'name': 451,
 'either': 452,
 'turn': 453,
 'sex': 454,
 'turns': 455,
 'care': 456,
 'one.': 457,
 'starts': 458,
 'written': 459,
 'dead': 460,
 'problem': 461,
 'episode': 462,
 "won't": 463,
 'this.': 464,
 'low': 465,
 'school': 466,
 'stupid': 467,
 'mean': 468,
 'face': 469,
 'behind': 470,
 'this,': 471,
 'me.': 472,
 'him.': 473,
 'moments': 474,
 'lead': 475,
 'favorite': 476,
 'and,': 477,
 'lines': 478,
 'sound': 479,
 'white': 480,
 'head': 481,
 'michael': 482,
 'stars': 483,
```

```
'guess': 484,
'took': 485,
'cannot': 486,
'story.': 487,
'fine': 488,
'budget': 489,
'title': 490,
'good.': 491,
'well.': 492,
'me,': 493,
'kill': 494,
'terrible': 495,
'highly': 496,
'extremely': 497,
'others': 498,
'all.': 499,
'dialogue': 500,
'sometimes': 501,
"film's": 502,
'evil': 503,
'beginning': 504,
'lack': 505,
'life.': 506,
'style': 507,
'throughout': 508,
"wouldn't": 509,
'heard': 510,
'obviously': 511,
'boring': 512,
'dark': 513,
'feeling': 514,
'itself': 515,
'fans': 516,
'lives': 517,
'decent': 518,
'works': 519,
'expect': 520,
'soon': 521,
'killer': 522,
'boy': 523,
'good,': 524,
'looked': 525,
'mr.': 526,
'attempt': 527,
'late': 528,
'particularly': 529,
'case': 530,
'amazing': 531,
'course,': 532,
'directed': 533,
'friend': 534,
'fight': 535,
'leave': 536,
'/>in': 537,
'wrong': 538,
'quality': 539,
'/>it': 540,
'mother': 541,
'picture': 542,
'entertaining': 543,
'save': 544,
```

```
 'complete': 545,
 'thinking': 546,
 'close': 547,
 'viewer': 548,
 'exactly': 549,
 'guys': 550,
 'taken': 551,
 'awful': 552,
 'writing': 553,
 'finds': 554,
 'run': 555,
 'whose': 556,
 'except': 557,
 '/>if': 558,
 'wonder': 559,
 'movies.': 560,
 'somewhat': 561,
 'living': 562,
 'across': 563,
 'number': 564,
 'police': 565,
 'strong': 566,
 'told': 567,
 'says': 568,
 'james': 569,
 'movies,': 570,
 'opening': 571,
 'worse': 572,
 'obvious': 573,
 'shown': 574,
 'parts': 575,
 'laugh': 576,
 'coming': 577,
 'running': 578,
 '3': 579,
 'car': 580,
 'films,': 581,
 'past': 582,
 'direction': 583,
 'usually': 584,
 'type': 585,
 'wish': 586,
 'group': 587,
 'huge': 588,
 'side': 589,
 'acting,': 590,
 'major': 591,
 'taking': 592,
 'children': 593,
 'supporting': 594,
 'girls': 595,
 'stop': 596,
 'act': 597,
 'bad.': 598,
 'hour': 599,
 'turned': 600,
 'tells': 601,
 'known': 602,
 'none': 603,
 'started': 604,
 'local': 605,
```

```
'knew': 606,
'that.': 607,
'myself': 608,
'killed': 609,
'due': 610,
'here,': 611,
'stories': 612,
"aren't": 613,
'town': 614,
'single': 615,
'game': 616,
'happens': 617,
',': 618,
'female': 619,
'it.<br': 620,
'again,': 621,
'bring': 622,
'call': 623,
'including': 624,
"i'll": 625,
'way,': 626,
'brilliant': 627,
'cinema': 628,
'here.': 629,
'son': 630,
'clearly': 631,
'british': 632,
'actress': 633,
'fact,': 634,
'/>there': 635,
'mostly': 636,
'humor': 637,
'art': 638,
'david': 639,
'hit': 640,
'talking': 641,
'voice': 642,
'robert': 643,
'involved': 644,
'whether': 645,
'characters,': 646,
'giving': 647,
'child': 648,
'musical': 649,
'one,': 650,
'history': 651,
'order': 652,
'out.': 653,
'ends': 654,
'falls': 655,
'end,': 656,
'heart': 657,
'saying': 658,
'drama': 659,
'easily': 660,
'!': 661,
'matter': 662,
'serious': 663,
'again.': 664,
'bad,': 665,
'eyes': 666,
```

```
'relationship': 667,
'horrible': 668,
'yes,': 669,
'nearly': 670,
'however': 671,
'modern': 672,
'chance': 673,
'feels': 674,
'beyond': 675,
'kid': 676,
'needs': 677,
'actual': 678,
'cut': 679,
"haven't": 680,
'appears': 681,
'important': 682,
'similar': 683,
'themselves': 684,
'five': 685,
'example': 686,
'moment': 687,
'/>but': 688,
'change': 689,
'comic': 690,
'upon': 691,
'end.': 692,
'simple': 693,
'strange': 694,
'using': 695,
'four': 696,
'knows': 697,
'happened': 698,
'is.': 699,
'named': 700,
'within': 701,
'released': 702,
'usual': 703,
'way.': 704,
'seen.': 705,
'english': 706,
'among': 707,
'movie.<br': 708,
'but,': 709,
'lots': 710,
'/>': 711,
'basically': 712,
'mention': 713,
'slow': 714,
'romantic': 715,
'stuff': 716,
'bunch': 717,
'her.': 718,
'kept': 719,
'typical': 720,
'certain': 721,
'interest': 722,
'films.': 723,
'near': 724,
'city': 725,
'hours': 726,
'overall': 727,
```

```
'showing': 728,
'(i': 729,
'blood': 730,
'plot,': 731,
'brought': 732,
'fall': 733,
'murder': 734,
'days': 735,
'body': 736,
'/>and': 737,
'tried': 738,
'song': 739,
'apparently': 740,
'miss': 741,
'/>a': 742,
'him,': 743,
'middle': 744,
'greatest': 745,
'jack': 746,
'life,': 747,
'george': 748,
'film.<br': 749,
'score': 750,
'events': 751,
'happy': 752,
'sad': 753,
'/>as': 754,
'yourself': 755,
'working': 756,
'shots': 757,
'stay': 758,
'talk': 759,
'add': 760,
'cheap': 761,
'(as': 762,
'begins': 763,
'decided': 764,
'age': 765,
'buy': 766,
'so,': 767,
'surprised': 768,
'hear': 769,
'french': 770,
'brother': 771,
'famous': 772,
'alone': 773,
'on.': 774,
'paul': 775,
'became': 776,
"what's": 777,
'daughter': 778,
'hate': 779,
'annoying': 780,
'learn': 781,
'richard': 782,
'happen': 783,
'(which': 784,
'please': 785,
'sit': 786,
'songs': 787,
"you've": 788,
```

```
 'de': 789,
 'jokes': 790,
 'easy': 791,
 "/>it's": 792,
 'nor': 793,
 'ten': 794,
 'peter': 795,
 'sets': 796,
 'documentary': 797,
 'view': 798,
 'experience': 799,
 'light': 800,
 'too.': 801,
 'above': 802,
 'funny.': 803,
 'cinematography': 804,
 'funny,': 805,
 'hell': 806,
 'possibly': 807,
 "who's": 808,
 'sequence': 809,
 'sexual': 810,
 'attention': 811,
 'king': 812,
 'straight': 813,
 '1': 814,
 'stand': 815,
 'hand': 816,
 'meets': 817,
 'power': 818,
 'oh': 819,
 'gore': 820,
 'episodes': 821,
 'word': 822,
 'difficult': 823,
 'clear': 824,
 'elements': 825,
 'husband': 826,
 'also,': 827,
 'red': 828,
 'god': 829,
 'violence': 830,
 'characters.': 831,
 'keeps': 832,
 'leaves': 833,
 'poorly': 834,
 'realize': 835,
 'means': 836,
 'ones': 837,
 'japanese': 838,
 'television': 839,
 'genre': 840,
 'roles': 841,
 'character,': 842,
 'somehow': 843,
 'flick': 844,
 '5': 845,
 'cool': 846,
 'out,': 847,
 'figure': 848,
 'reality': 849,
```

```
'possible': 850,
'silly': 851,
'them,': 852,
'eventually': 853,
'towards': 854,
'reading': 855,
'move': 856,
'rent': 857,
'brings': 858,
'doubt': 859,
'scary': 860,
'theme': 861,
'killing': 862,
'lady': 863,
'unfortunately': 864,
'filmed': 865,
'imagine': 866,
'moving': 867,
'previous': 868,
'emotional': 869,
'on,': 870,
'leads': 871,
'rock': 872,
'room': 873,
'(or': 874,
'problems': 875,
'third': 876,
'tom': 877,
'deal': 878,
'say,': 879,
'oscar': 880,
'though,': 881,
'review': 882,
'talent': 883,
'novel': 884,
'comments': 885,
'forget': 886,
'hero': 887,
'leading': 888,
'message': 889,
'enjoyable': 890,
'incredibly': 891,
'level': 892,
'write': 893,
'career': 894,
'ridiculous': 895,
'check': 896,
'better.': 897,
'various': 898,
'plenty': 899,
'animation': 900,
'meant': 901,
'needed': 902,
'personal': 903,
'feature': 904,
'show.': 905,
'writer': 906,
'future': 907,
'herself': 908,
'avoid': 909,
'gone': 910,
```

```
'meet': 911,
'create': 912,
'work.': 913,
'hardly': 914,
'begin': 915,
'whom': 916,
'total': 917,
'country': 918,
'manages': 919,
'team': 920,
'theater': 921,
'watch.': 922,
'4': 923,
'forced': 924,
'fairly': 925,
'form': 926,
'up.': 927,
'particular': 928,
'there.': 929,
'hilarious': 930,
'appear': 931,
'points': 932,
'you.': 933,
'now,': 934,
'viewers': 935,
'fantastic': 936,
'effort': 937,
'interested': 938,
'follow': 939,
'unless': 940,
'weak': 941,
'words': 942,
'fast': 943,
'dramatic': 944,
'expecting': 945,
'tale': 946,
'unfortunately,': 947,
'joe': 948,
'zombie': 949,
'male': 950,
'older': 951,
'average': 952,
'her,': 953,
'general': 954,
'political': 955,
'attempts': 956,
'reviews': 957,
'mystery': 958,
'scenes,': 959,
'features': 960,
'sounds': 961,
'subject': 962,
'(a': 963,
'class': 964,
'whatever': 965,
'york': 966,
'pay': 967,
'crap': 968,
'lee': 969,
'up,': 970,
'scene,': 971,
```

```
                'worked': 972,
                'dance': 973,
                'decides': 974,
                'disney': 975,
                'premise': 976,
                'then,': 977,
                'plain': 978,
                'man,': 979,
                "let's": 980,
                'minute': 981,
                'badly': 982,
                'ask': 983,
                'dialog': 984,
                '20': 985,
                'expected': 986,
                'viewing': 987,
                'forward': 988,
                'open': 989,
                'uses': 990,
                'wait': 991,
                '(who': 992,
                'deep': 993,
                'social': 994,
                'storyline': 995,
                'front': 996,
                'crime': 997,
                'directors': 998,
                'writers': 999,
                ...})
```

In [24]:

```
#Split the trainset into trainset and validation set
trainset, valset = trainset.split(split_ratio=0.8)
```

In [25]:

```
#Make the data loader
train_iter, val_iter, test_iter = torchtext.data.BucketIterator.splits(
        (trainset, valset, testset), batch_size=64,
        shuffle=True, repeat=False)
```

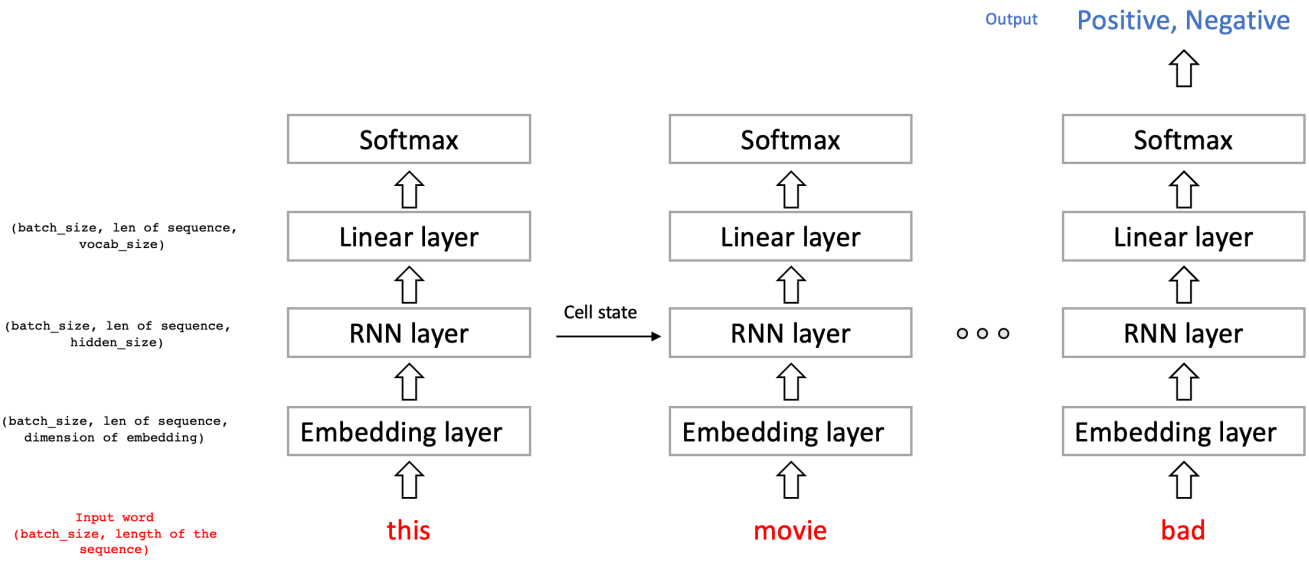
In [26]:

```
#Since the length of the sentence is different, the size of the data varies from batch to batch
batch = next(iter(train_iter))
print(f'Shape of first batch: {batch.text.shape}')
batch = next(iter(train_iter))
print(f'Shape of second batch: {batch.text.shape}')

#Reset the data loader to learn the observed data as well
train_iter, val_iter, test_iter =torchtext.data.BucketIterator.splits(
        (trainset, valset, testset), batch_size=64,
        shuffle=True, repeat=False)
```

```
Shape of first batch: torch.Size([64, 967])
Shape of second batch: torch.Size([64, 863])
```

2-2. Implenment the RNN model



In [27]:

```
class classification_RNN(torch.nn.Module):
  def __init__(self, vocab_size, input_size, n_labels, hidden_size, batch_first=True):
    super(classification_RNN, self).__init__()
    self.vocab_size = vocab_size
    self.input_size = input_size
    self.hidden_size = hidden_size

    #Embedding layer
    self.embedding_layer = torch.nn.Embedding(num_embeddings=vocab_size,embedding_dim=input_size
)

    #RNN layer
    self.rnn_layer = torch.nn.RNN(input_size=input_size, hidden_size=hidden_size, batch_first=ba
tch_first)

    #linear layer
    self.linear = torch.nn.Linear(hidden_size, n_labels)


  def forward(self, x):
    #1. Embedding layer
    #size of the data: (batch_size, length of the sequence)
    #-> (batch_size, length of the sequence, dimension of embedding)
    y = self.embedding_layer(x)

    #2. RNN layer
    #size of the data: (batch_size, length of the sequence, dimension of embedding)
    #-> y: (batch_size, lenth of the sequence, hidden_size), hidden: (1, batch_size, hidden_siz
e)
    y, hidden = self.rnn_layer(y)

    #Using only last time step result
    #y = (batch_size, hidden_size)
    y = y[:,-1,:]

    #3. Linear layer
    #size of the data: (batch_size, hidden_size)
    #-> (batch_size, n_labels)
    y = self.linear(y)

    #Size of the return value: (batch_size*length of the sequence, vocab_size)
    return y
```

2-3. Train the RNN model

In [28]:

```
#hyper parameter
lr = 0.001
epochs = 10
```

In [29]:

```
model = classification_RNN(vocab_size=vocab_size, input_size = 128, n_labels=2, hidden_size=256)
.to(DEVICE)
optimizer = torch.optim.Adam(model.parameters(), lr=lr)
```

In [30]:

```
#train function
def train(model, optimizer, train_iter):
  for b, batch in enumerate(train_iter):
    x, y = batch.text.to(DEVICE), batch.label.to(DEVICE)
    y.data.sub_(1)  #transforms the labels into (0, 1)
    optimizer.zero_grad()

    logit = model(x)
    loss = torch.nn.functional.cross_entropy(logit, y)
    loss.backward()
    optimizer.step()
```

In [31]:

```
#Evaluate function
def evaluate(model, val_iter):
  corrects, total_loss = 0, 0
  for batch in val_iter:
    x, y = batch.text.to(DEVICE), batch.label.to(DEVICE)
    y.data.sub_(1) #transforms the labels into (0, 1)
    logit = model(x)
    loss = torch.nn.functional.cross_entropy(logit, y, reduction='sum')
    total_loss += loss.item()
    corrects += (logit.max(1)[1].view(y.size()).data == y.data).sum()
  size = len(val_iter.dataset)
  avg_loss = total_loss / size
  avg_accuracy = 100.0 * corrects / size
  return avg_loss, avg_accuracy
```

In [32]:

```
for e in range(1, epochs+1):
    train(model, optimizer, train_iter)
    val_loss, val_accuracy = evaluate(model, val_iter)
    print(f"Epoch: {e}, Loss of validation: {val_loss} Accuracy of validation: {val_accuracy}")
```

```
Epoch: 1, Loss of validation: 0.6988149040222168 Accuracy of validation: 49.5
Epoch: 2, Loss of validation: 0.6964097266197204 Accuracy of validation: 50.520000
45776367
Epoch: 3, Loss of validation: 0.7064652287483215 Accuracy of validation: 49.579998
01635742
Epoch: 4, Loss of validation: 0.6935130256652832 Accuracy of validation: 52.039997
10083008
Epoch: 5, Loss of validation: 0.6932604013442993 Accuracy of validation: 52.020000
45776367
Epoch: 6, Loss of validation: 0.6938436466217041 Accuracy of validation: 52.059997
55859375
Epoch: 7, Loss of validation: 0.696732887172699 Accuracy of validation: 48.5599975
5859375
Epoch: 8, Loss of validation: 0.7038828132629394 Accuracy of validation: 47.840000
15258789
Epoch: 9, Loss of validation: 0.7061405626296997 Accuracy of validation: 48.979999
54223633
Epoch: 10, Loss of validation: 0.6953793263435364 Accuracy of validation: 51.0
```

In [33]:

```
test_loss, test_acc = evaluate(model, test_iter)
print(f'Test loss: {test_loss}, Test accuracy: {test_acc}')
```

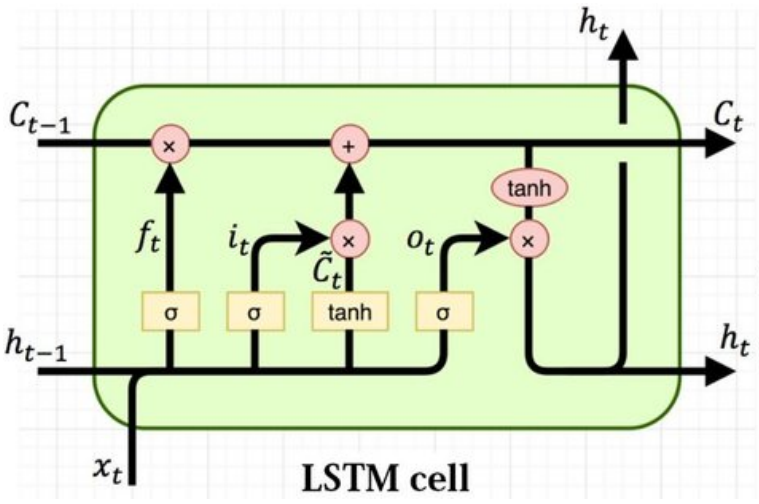
```
Test loss: 0.6749759851837158, Test accuracy: 62.279998779296875
```

Problem 1(10pt)

Implement LSTM model and Train that model

*The test accuracy of your lstm model should be higher than that of the rnn model.

*Fill the ### blank of below model.



$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right)$$
$$f_t = \sigma\left(x_t U^f + h_{t-1} W^f\right)$$
$$o_t = \sigma\left(x_t U^o + h_{t-1} W^o\right)$$
$$\tilde{C}_t = \tanh\left(x_t U^g + h_{t-1} W^g\right)$$
$$C_t = \sigma\left(f_t * C_{t-1} + i_t * \tilde{C}_t\right)$$
$$h_t = \tanh(C_t) * o_t$$

In [34]:

```
##https://pytorch.org/docs/stable/generated/torch.nn.LSTM.html
##https://stackoverflow.com/questions/48302810/whats-the-difference-between-hidden-and-output-in-pytorch-lstm


class LSTM(torch.nn.Module):
  def __init__(self, vocab_size, input_size, n_labels, hidden_size, num_layers = 1, batch_first=True):
    super(LSTM, self).__init__()
    self.vocab_size = vocab_size
    self.input_size = input_size
    self.hidden_size = hidden_size
    self.num_layers = num_layers
    #Embedding layer
    self.embedding_layer = torch.nn.Embedding(num_embeddings=vocab_size,embedding_dim=input_size)

    #RNN layer
    self.lstm_layer = torch.nn.LSTM(input_size=input_size, hidden_size=hidden_size, num_layers = num_layers,batch_first=batch_first)

    #linear layer
    self.linear = torch.nn.Linear(hidden_size, vocab_size)

  def forward(self, x):
    #1. Embedding layer
    #size of the data: (batch_size, lenth of the sequence)
    #-> (batch_size, lenth of the sequence, dimension of embedding)
    y = self.embedding_layer(x)

    #Initial hidden state
    h_0 = torch.zeros((self.num_layers, y.shape[0], self.hidden_size)).to(DEVICE)

    #Initial cell state
    c_0 = torch.zeros((self.num_layers, y.shape[0], self.hidden_size)).to(DEVICE)

    #2. LSTM layer
    #size of the data: (batch_size, lenth of the sequence, dimension of embedding)
    #-> h_n, c_0 = (num_layers, batch size, hidden_size)
    hidden_states, (h_n , c_n) = self.lstm_layer(y, (h_0, c_0))

    #h_n = (num_layers*batch size, hidden_size)
    h_n = h_n.view(h_n.shape[1] ,-1)

    #3. Linear layer
    #size of the data: (batch_size, hidden_size)
    #-> (batch_size, n_labels)
    result = self.linear(h_n)

    #Size of the return value: (batch_size, n_labels)
    return result
```

In [35]:

```
model = LSTM(vocab_size=vocab_size, input_size=128, n_labels=2, hidden_size=256).to(DEVICE)
optimizer = torch.optim.Adam(model.parameters(), lr=lr)
```

In [36]:

```
for e in range(1, epochs+1):
    train(model, optimizer, train_iter)
    val_loss, val_accuracy = evaluate(model, val_iter)
    print(f"Epoch: {e}, Loss of validation: {val_loss} Accuracy of validation: {val_accuracy}")
```

```
Epoch: 1, Loss of validation: 0.6961421512603759 Accuracy of validation: 51.259998
3215332
Epoch: 2, Loss of validation: 0.7177148748397827 Accuracy of validation: 51.259998
3215332
Epoch: 3, Loss of validation: 0.7136506860733032 Accuracy of validation: 51.259998
3215332
Epoch: 4, Loss of validation: 0.6971550874710083 Accuracy of validation: 50.379997
25341797
Epoch: 5, Loss of validation: 0.6999296840667725 Accuracy of validation: 51.259998
3215332
Epoch: 6, Loss of validation: 0.697055998802185 Accuracy of validation: 50.2599983
215332
Epoch: 7, Loss of validation: 0.7167784940719605 Accuracy of validation: 48.719997
40600586
Epoch: 8, Loss of validation: 0.6967366484642029 Accuracy of validation: 50.479999
54223633
Epoch: 9, Loss of validation: 0.7020181018829346 Accuracy of validation: 50.939998
626708984
Epoch: 10, Loss of validation: 0.7036067196846009 Accuracy of validation: 48.59999
8474121094
```

In [37]:

```
test_loss, test_acc = evaluate(model, test_iter)
print(f'Test loss: {test_loss}, Test accuracy: {test_acc}')
```

```
Test loss: 0.6890623641967774, Test accuracy: 48.01599884033203
```

Problem 2(10 points)

Explain why the LSTM model performs better than simple RNN model.

Write your answer

In [38]:

```
# Rnn은 역전파시 그 거리가 멀경우 그래디언트가 줄어들기 때문에 학습이 잘 안되지만 Lstm의 경우 ce
ll state을 추가하기때문에 이 문제를 없애줍니다.
```

Problem 3(10 points)

Implement the LSTM model instead of RNN in problem1(many to many problem) above

In [39]:

```
#Write the sentence
sentence = "Eagle Rare is a bourbon whisky that i like the most".split()

#Make the vokabulary
vocab = list(set(sentence))
word2index = {tkn: i for i, tkn in enumerate(vocab, 1)}
word2index['<unk>']=0
index2word = {v: k for k, v in word2index.items()}
x, y, encoded = build_data(sentence, word2index)
embedding_function = torch.nn.Embedding(num_embeddings=len(word2index), embedding_dim = 5)
embedding_function(x)
```

Out[39]:

```
tensor([[[ 1.3344, -0.4350,  1.4764, -0.2917, -1.3991],
         [-1.3174,  0.1444,  0.0370, -1.6321, -0.2543],
         [ 0.8664,  0.0514,  0.9300, -1.3784,  0.1689],
         [-0.6535, -0.4162, -0.8643,  0.3408, -0.6381],
         [ 0.1855,  0.4339, -0.4353,  0.8423, -0.0186],
         [-0.0052, -0.0983, -0.4413,  2.6835, -1.0700],
         [ 0.3426,  0.8962,  1.5003, -0.2158, -0.2627],
         [-0.5942,  2.1728,  1.6469, -1.7226,  0.0259],
         [ 1.7063,  2.6437,  0.5030, -0.4673, -0.5152],
         [-0.1492, -0.6354, -0.1864, -0.0843,  0.9663]]],
       grad_fn=<EmbeddingBackward>)
```

In [40]:

```
# implement the LSTM model

class LSTM_model(torch.nn.Module):
  def __init__(self, vocab_size, input_size, hidden_size, num_layers = 1, batch_first=True):
    super(LSTM_model, self).__init__()
    self.vocab_size = vocab_size
    self.input_size = input_size
    self.hidden_size = hidden_size
    self.num_layers = num_layers
    #Embedding layer
    self.embedding_layer = torch.nn.Embedding(num_embeddings=vocab_size,embedding_dim=input_size
)

    #RNN layer
    self.lstm_layer = torch.nn.LSTM(input_size=input_size, hidden_size=hidden_size, num_layers =
num_layers,batch_first=batch_first)

    #linear layer
    self.linear = torch.nn.Linear(hidden_size, vocab_size)

  def forward(self, x):
    #1. Embedding layer
    #size of the data: (batch_size, lenth of the sequence)
    #-> (batch_size, lenth of the sequence, dimension of embedding)
    y = self.embedding_layer(x)

    #Initial hidden state
    h_0 = torch.zeros((self.num_layers, y.shape[0], self.hidden_size)).to(DEVICE)

    #Initial cell state
    c_0 = torch.zeros((self.num_layers, y.shape[0], self.hidden_size)).to(DEVICE)

    #2. LSTM layer
    #size of the data: (batch_size, lenth of the sequence, dimension of embedding)
    #-> h_n, c_0 = (num_layers, batch size, hidden_size)
    hidden_states, (h_n , c_n) = self.lstm_layer(y, (h_0, c_0))

    #h_n = (num_layers*batch size, hidden_size)
    #h_n = h_n.view(h_n.shape[1] ,-1)

    #3. Linear layer
    #size of the data: (batch_size, hidden_size)
    #-> (batch_size, n_labels)
    #result = self.linear(h_n)
    result = self.linear(hidden_states)

    #Size of the return value: (batch_size, n_labels)
    return result.view(-1, result.size(2))
```

In [41]:

```
#hyper parameter
vocab_size = len(word2index)
input_size = 5
hidden_size = 20
epochs = 100

model = LSTM_model(vocab_size=vocab_size, input_size=input_size, hidden_size=hidden_size).to(DEV
ICE)
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(params=model.parameters())

#Training
for epoch in range(1, epochs+1):
    optimizer.zero_grad()
    output = model(x.to(DEVICE))
    loss = loss_function(output, y.view(-1).to(DEVICE))
    loss.backward()
    optimizer.step()
    # Observe the result
    if epoch % 20 == 0:
        print(f"Epoch: {epoch}, Loss: {loss}")
        pred = output.softmax(-1).argmax(-1).tolist()
        print(" ".join(['All']+ decode(pred)))
        print()
```

```
Epoch: 20, Loss: 2.4139750003814697
All bourbon bourbon a bourbon bourbon bourbon like like like a

Epoch: 40, Loss: 2.3315021991729736
All a bourbon a bourbon bourbon bourbon like like like a

Epoch: 60, Loss: 2.1930038928985596
All a a a bourbon bourbon bourbon like like the most

Epoch: 80, Loss: 1.9410358667373657
All Rare a a bourbon bourbon that i the the most

Epoch: 100, Loss: 1.6363279819488525
All Rare a a bourbon bourbon that i the the most
```

In [ ]:

```
!apt-get install texlive texlive-xetex texlive-latex-extra pandoc
!pip install pypandoc
from google.colab import drive
drive.mount('/content/drive')
!jupyter nbconvert --to html '/content/drive/MyDrive/인공지능프로젝트/swe3032_week6.ipynb'
```

```
Reading package lists... Done
Building dependency tree
Reading state information... Done
pandoc is already the newest version (1.19.2.4~dfsg-1build4).
pandoc set to manually installed.
The following package was automatically installed and is no longer required:
  libnvidia-common-460
Use 'apt autoremove' to remove it.
The following additional packages will be installed:
  fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono fonts-texgyre
  javascript-common libcupsfilters1 libcupsimage2 libgs9 libgs9-common
  libijs-0.35 libjbig2dec0 libjs-jquery libkpathsea6 libpotrace0 libptexenc1
  libruby2.5 libsynctex1 libtexlua52 libtexluajit2 libzzip-0-13 lmodern
  poppler-data preview-latex-style rake ruby ruby-did-you-mean ruby-minitest
  ruby-net-telnet ruby-power-assert ruby-test-unit ruby2.5
  rubygems-integration t1utils tex-common tex-gyre texlive-base
  texlive-binaries texlive-fonts-recommended texlive-latex-base
  texlive-latex-recommended texlive-pictures texlive-plain-generic tipa
Suggested packages:
  fonts-noto apache2 | lighttpd | httpd poppler-utils ghostscript
  fonts-japanese-mincho | fonts-ipafont-mincho fonts-japanese-gothic
  | fonts-ipafont-gothic fonts-arphic-ukai fonts-arphic-uming fonts-nanum ri
  ruby-dev bundler debhelper gv | postscript-viewer perl-tk xpdf-reader
  | pdf-viewer texlive-fonts-recommended-doc texlive-latex-base-doc
  python-pygments icc-profiles libfile-which-perl
  libspreadsheet-parseexcel-perl texlive-latex-extra-doc
  texlive-latex-recommended-doc texlive-pstricks dot2tex prerex ruby-tcltk
  | libtcltk-ruby texlive-pictures-doc vprerex
The following NEW packages will be installed:
  fonts-droid-fallback fonts-lato fonts-lmodern fonts-noto-mono fonts-texgyre
  javascript-common libcupsfilters1 libcupsimage2 libgs9 libgs9-common
  libijs-0.35 libjbig2dec0 libjs-jquery libkpathsea6 libpotrace0 libptexenc1
  libruby2.5 libsynctex1 libtexlua52 libtexluajit2 libzzip-0-13 lmodern
  poppler-data preview-latex-style rake ruby ruby-did-you-mean ruby-minitest
  ruby-net-telnet ruby-power-assert ruby-test-unit ruby2.5
  rubygems-integration t1utils tex-common tex-gyre texlive texlive-base
  texlive-binaries texlive-fonts-recommended texlive-latex-base
  texlive-latex-extra texlive-latex-recommended texlive-pictures
  texlive-plain-generic texlive-xetex tipa
0 upgraded, 47 newly installed, 0 to remove and 12 not upgraded.
Need to get 146 MB of archives.
After this operation, 460 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu bionic/main amd64 fonts-droid-fallback all
1:6.0.1r16-1.1 [1,805 kB]
Get:2 http://archive.ubuntu.com/ubuntu bionic/main amd64 fonts-lato all 2.0-2 [2,6
98 kB]
Get:3 http://archive.ubuntu.com/ubuntu bionic/main amd64 poppler-data all 0.4.8-2
[1,479 kB]
Get:4 http://archive.ubuntu.com/ubuntu bionic/main amd64 tex-common all 6.09 [33.0
kB]
Get:5 http://archive.ubuntu.com/ubuntu bionic/main amd64 fonts-lmodern all 2.004.5
-3 [4,551 kB]
Get:6 http://archive.ubuntu.com/ubuntu bionic/main amd64 fonts-noto-mono all 20171
026-2 [75.5 kB]
Get:7 http://archive.ubuntu.com/ubuntu bionic/universe amd64 fonts-texgyre all 201
60520-1 [8,761 kB]
Get:8 http://archive.ubuntu.com/ubuntu bionic/main amd64 javascript-common all 11
[6,066 B]
Get:9 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libcupsfilters1 a
md64 1.20.2-0ubuntu3.1 [108 kB]
Get:10 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libcupsimage2 am
```

```
d64 2.2.7-1ubuntu2.9 [18.6 kB]
Get:11 http://archive.ubuntu.com/ubuntu bionic/main amd64 libijs-0.35 amd64 0.35-1
3 [15.5 kB]
Get:12 http://archive.ubuntu.com/ubuntu bionic/main amd64 libjbig2dec0 amd64 0.13-
6 [55.9 kB]
Get:13 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libgs9-common al
l 9.26~dfsg+0-0ubuntu0.18.04.17 [5,092 kB]
Get:14 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libgs9 amd64 9.2
6~dfsg+0-0ubuntu0.18.04.17 [2,267 kB]
Get:15 http://archive.ubuntu.com/ubuntu bionic/main amd64 libjs-jquery all 3.2.1-1
[152 kB]
Get:16 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libkpathsea6 amd
64 2017.20170613.44572-8ubuntu0.1 [54.9 kB]
Get:17 http://archive.ubuntu.com/ubuntu bionic/main amd64 libpotrace0 amd64 1.14-2
[17.4 kB]
Get:18 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libptexenc1 amd6
4 2017.20170613.44572-8ubuntu0.1 [34.5 kB]
Get:19 http://archive.ubuntu.com/ubuntu bionic/main amd64 rubygems-integration all
1.11 [4,994 B]
Get:20 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 ruby2.5 amd64 2.
5.1-1ubuntu1.12 [48.6 kB]
Get:21 http://archive.ubuntu.com/ubuntu bionic/main amd64 ruby amd64 1:2.5.1 [5,71
2 B]
Get:22 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 rake all 12.3.1-
1ubuntu0.1 [44.9 kB]
Get:23 http://archive.ubuntu.com/ubuntu bionic/main amd64 ruby-did-you-mean all 1.
2.0-2 [9,700 B]
Get:24 http://archive.ubuntu.com/ubuntu bionic/main amd64 ruby-minitest all 5.10.3
-1 [38.6 kB]
Get:25 http://archive.ubuntu.com/ubuntu bionic/main amd64 ruby-net-telnet all 0.1.
1-2 [12.6 kB]
Get:26 http://archive.ubuntu.com/ubuntu bionic/main amd64 ruby-power-assert all 0.
3.0-1 [7,952 B]
Get:27 http://archive.ubuntu.com/ubuntu bionic/main amd64 ruby-test-unit all 3.2.5
-1 [61.1 kB]
Get:28 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libruby2.5 amd64
2.5.1-1ubuntu1.12 [3,073 kB]
Get:29 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libsynctex1 amd6
4 2017.20170613.44572-8ubuntu0.1 [41.4 kB]
Get:30 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libtexlua52 amd6
4 2017.20170613.44572-8ubuntu0.1 [91.2 kB]
Get:31 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libtexluajit2 am
d64 2017.20170613.44572-8ubuntu0.1 [230 kB]
Get:32 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 libzzip-0-13 amd
64 0.13.62-3.1ubuntu0.18.04.1 [26.0 kB]
Get:33 http://archive.ubuntu.com/ubuntu bionic/main amd64 lmodern all 2.004.5-3
[9,631 kB]
Get:34 http://archive.ubuntu.com/ubuntu bionic/main amd64 preview-latex-style all
11.91-1ubuntu1 [185 kB]
Get:35 http://archive.ubuntu.com/ubuntu bionic/main amd64 t1utils amd64 1.41-2 [5
6.0 kB]
Get:36 http://archive.ubuntu.com/ubuntu bionic/universe amd64 tex-gyre all 2016052
0-1 [4,998 kB]
Get:37 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 texlive-binaries
amd64 2017.20170613.44572-8ubuntu0.1 [8,179 kB]
Get:38 http://archive.ubuntu.com/ubuntu bionic/main amd64 texlive-base all 2017.20
180305-1 [18.7 MB]
Get:39 http://archive.ubuntu.com/ubuntu bionic/universe amd64 texlive-fonts-recomm
ended all 2017.20180305-1 [5,262 kB]
Get:40 http://archive.ubuntu.com/ubuntu bionic/main amd64 texlive-latex-base all 2
017.20180305-1 [951 kB]
```

```
Get:41 http://archive.ubuntu.com/ubuntu bionic/main amd64 texlive-latex-recommende
d all 2017.20180305-1 [14.9 MB]
Get:42 http://archive.ubuntu.com/ubuntu bionic/universe amd64 texlive all 2017.201
80305-1 [14.4 kB]
Get:43 http://archive.ubuntu.com/ubuntu bionic/universe amd64 texlive-pictures all
2017.20180305-1 [4,026 kB]
Get:44 http://archive.ubuntu.com/ubuntu bionic/universe amd64 texlive-latex-extra
all 2017.20180305-2 [10.6 MB]
Get:45 http://archive.ubuntu.com/ubuntu bionic/universe amd64 texlive-plain-generi
c all 2017.20180305-2 [23.6 MB]
Get:46 http://archive.ubuntu.com/ubuntu bionic/universe amd64 tipa all 2:1.3-20
[2,978 kB]
Get:47 http://archive.ubuntu.com/ubuntu bionic/universe amd64 texlive-xetex all 20
17.20180305-1 [10.7 MB]
Fetched 146 MB in 11s (13.2 MB/s)
Extracting templates from packages: 100%
Preconfiguring packages ...
Selecting previously unselected package fonts-droid-fallback.
(Reading database ... 123934 files and directories currently installed.)
Preparing to unpack .../00-fonts-droid-fallback_1%3a6.0.1r16-1.1_all.deb ...
Unpacking fonts-droid-fallback (1:6.0.1r16-1.1) ...
Selecting previously unselected package fonts-lato.
Preparing to unpack .../01-fonts-lato_2.0-2_all.deb ...
Unpacking fonts-lato (2.0-2) ...
Selecting previously unselected package poppler-data.
Preparing to unpack .../02-poppler-data_0.4.8-2_all.deb ...
Unpacking poppler-data (0.4.8-2) ...
Selecting previously unselected package tex-common.
Preparing to unpack .../03-tex-common_6.09_all.deb ...
Unpacking tex-common (6.09) ...
Selecting previously unselected package fonts-lmodern.
Preparing to unpack .../04-fonts-lmodern_2.004.5-3_all.deb ...
Unpacking fonts-lmodern (2.004.5-3) ...
Selecting previously unselected package fonts-noto-mono.
Preparing to unpack .../05-fonts-noto-mono_20171026-2_all.deb ...
Unpacking fonts-noto-mono (20171026-2) ...
Selecting previously unselected package fonts-texgyre.
Preparing to unpack .../06-fonts-texgyre_20160520-1_all.deb ...
Unpacking fonts-texgyre (20160520-1) ...
Selecting previously unselected package javascript-common.
Preparing to unpack .../07-javascript-common_11_all.deb ...
Unpacking javascript-common (11) ...
Selecting previously unselected package libcupsfilters1:amd64.
Preparing to unpack .../08-libcupsfilters1_1.20.2-0ubuntu3.1_amd64.deb ...
Unpacking libcupsfilters1:amd64 (1.20.2-0ubuntu3.1) ...
Selecting previously unselected package libcupsimage2:amd64.
Preparing to unpack .../09-libcupsimage2_2.2.7-1ubuntu2.9_amd64.deb ...
Unpacking libcupsimage2:amd64 (2.2.7-1ubuntu2.9) ...
Selecting previously unselected package libijs-0.35:amd64.
Preparing to unpack .../10-libijs-0.35_0.35-13_amd64.deb ...
Unpacking libijs-0.35:amd64 (0.35-13) ...
Selecting previously unselected package libjbig2dec0:amd64.
Preparing to unpack .../11-libjbig2dec0_0.13-6_amd64.deb ...
Unpacking libjbig2dec0:amd64 (0.13-6) ...
Selecting previously unselected package libgs9-common.
Preparing to unpack .../12-libgs9-common_9.26~dfsg+0-0ubuntu0.18.04.17_all.deb ...
Unpacking libgs9-common (9.26~dfsg+0-0ubuntu0.18.04.17) ...
Selecting previously unselected package libgs9:amd64.
Preparing to unpack .../13-libgs9_9.26~dfsg+0-0ubuntu0.18.04.17_amd64.deb ...
Unpacking libgs9:amd64 (9.26~dfsg+0-0ubuntu0.18.04.17) ...
Selecting previously unselected package libjs-jquery.
```

```
Preparing to unpack .../14-libjs-jquery_3.2.1-1_all.deb ...
Unpacking libjs-jquery (3.2.1-1) ...
Selecting previously unselected package libkpathsea6:amd64.
Preparing to unpack .../15-libkpathsea6_2017.20170613.44572-8ubuntu0.1_amd64.deb
...
Unpacking libkpathsea6:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Selecting previously unselected package libpotrace0.
Preparing to unpack .../16-libpotrace0_1.14-2_amd64.deb ...
Unpacking libpotrace0 (1.14-2) ...
Selecting previously unselected package libptexenc1:amd64.
Preparing to unpack .../17-libptexenc1_2017.20170613.44572-8ubuntu0.1_amd64.deb
...
Unpacking libptexenc1:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Selecting previously unselected package rubygems-integration.
Preparing to unpack .../18-rubygems-integration_1.11_all.deb ...
Unpacking rubygems-integration (1.11) ...
Selecting previously unselected package ruby2.5.
Preparing to unpack .../19-ruby2.5_2.5.1-1ubuntu1.12_amd64.deb ...
Unpacking ruby2.5 (2.5.1-1ubuntu1.12) ...
Selecting previously unselected package ruby.
Preparing to unpack .../20-ruby_1%3a2.5.1_amd64.deb ...
Unpacking ruby (1:2.5.1) ...
Selecting previously unselected package rake.
Preparing to unpack .../21-rake_12.3.1-1ubuntu0.1_all.deb ...
Unpacking rake (12.3.1-1ubuntu0.1) ...
Selecting previously unselected package ruby-did-you-mean.
Preparing to unpack .../22-ruby-did-you-mean_1.2.0-2_all.deb ...
Unpacking ruby-did-you-mean (1.2.0-2) ...
Selecting previously unselected package ruby-minitest.
Preparing to unpack .../23-ruby-minitest_5.10.3-1_all.deb ...
Unpacking ruby-minitest (5.10.3-1) ...
Selecting previously unselected package ruby-net-telnet.
Preparing to unpack .../24-ruby-net-telnet_0.1.1-2_all.deb ...
Unpacking ruby-net-telnet (0.1.1-2) ...
Selecting previously unselected package ruby-power-assert.
Preparing to unpack .../25-ruby-power-assert_0.3.0-1_all.deb ...
Unpacking ruby-power-assert (0.3.0-1) ...
Selecting previously unselected package ruby-test-unit.
Preparing to unpack .../26-ruby-test-unit_3.2.5-1_all.deb ...
Unpacking ruby-test-unit (3.2.5-1) ...
Selecting previously unselected package libruby2.5:amd64.
Preparing to unpack .../27-libruby2.5_2.5.1-1ubuntu1.12_amd64.deb ...
Unpacking libruby2.5:amd64 (2.5.1-1ubuntu1.12) ...
Selecting previously unselected package libsynctex1:amd64.
Preparing to unpack .../28-libsynctex1_2017.20170613.44572-8ubuntu0.1_amd64.deb
...
Unpacking libsynctex1:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Selecting previously unselected package libtexlua52:amd64.
Preparing to unpack .../29-libtexlua52_2017.20170613.44572-8ubuntu0.1_amd64.deb
...
Unpacking libtexlua52:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Selecting previously unselected package libtexluajit2:amd64.
Preparing to unpack .../30-libtexluajit2_2017.20170613.44572-8ubuntu0.1_amd64.deb
...
Unpacking libtexluajit2:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Selecting previously unselected package libzzip-0-13:amd64.
Preparing to unpack .../31-libzzip-0-13_0.13.62-3.1ubuntu0.18.04.1_amd64.deb ...
Unpacking libzzip-0-13:amd64 (0.13.62-3.1ubuntu0.18.04.1) ...
Selecting previously unselected package lmodern.
Preparing to unpack .../32-lmodern_2.004.5-3_all.deb ...
Unpacking lmodern (2.004.5-3) ...
```

```
Selecting previously unselected package preview-latex-style.
Preparing to unpack .../33-preview-latex-style_11.91-1ubuntu1_all.deb ...
Unpacking preview-latex-style (11.91-1ubuntu1) ...
Selecting previously unselected package t1utils.
Preparing to unpack .../34-t1utils_1.41-2_amd64.deb ...
Unpacking t1utils (1.41-2) ...
Selecting previously unselected package tex-gyre.
Preparing to unpack .../35-tex-gyre_20160520-1_all.deb ...
Unpacking tex-gyre (20160520-1) ...
Selecting previously unselected package texlive-binaries.
Preparing to unpack .../36-texlive-binaries_2017.20170613.44572-8ubuntu0.1_amd64.d
eb ...
Unpacking texlive-binaries (2017.20170613.44572-8ubuntu0.1) ...
Selecting previously unselected package texlive-base.
Preparing to unpack .../37-texlive-base_2017.20180305-1_all.deb ...
Unpacking texlive-base (2017.20180305-1) ...
Selecting previously unselected package texlive-fonts-recommended.
Preparing to unpack .../38-texlive-fonts-recommended_2017.20180305-1_all.deb ...
Unpacking texlive-fonts-recommended (2017.20180305-1) ...
Selecting previously unselected package texlive-latex-base.
Preparing to unpack .../39-texlive-latex-base_2017.20180305-1_all.deb ...
Unpacking texlive-latex-base (2017.20180305-1) ...
Selecting previously unselected package texlive-latex-recommended.
Preparing to unpack .../40-texlive-latex-recommended_2017.20180305-1_all.deb ...
Unpacking texlive-latex-recommended (2017.20180305-1) ...
Selecting previously unselected package texlive.
Preparing to unpack .../41-texlive_2017.20180305-1_all.deb ...
Unpacking texlive (2017.20180305-1) ...
Selecting previously unselected package texlive-pictures.
Preparing to unpack .../42-texlive-pictures_2017.20180305-1_all.deb ...
Unpacking texlive-pictures (2017.20180305-1) ...
Selecting previously unselected package texlive-latex-extra.
Preparing to unpack .../43-texlive-latex-extra_2017.20180305-2_all.deb ...
Unpacking texlive-latex-extra (2017.20180305-2) ...
Selecting previously unselected package texlive-plain-generic.
Preparing to unpack .../44-texlive-plain-generic_2017.20180305-2_all.deb ...
Unpacking texlive-plain-generic (2017.20180305-2) ...
Selecting previously unselected package tipa.
Preparing to unpack .../45-tipa_2%3a1.3-20_all.deb ...
Unpacking tipa (2:1.3-20) ...
Selecting previously unselected package texlive-xetex.
Preparing to unpack .../46-texlive-xetex_2017.20180305-1_all.deb ...
Unpacking texlive-xetex (2017.20180305-1) ...
Setting up libgs9-common (9.26~dfsg+0-0ubuntu0.18.04.17) ...
Setting up libkpathsea6:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Setting up libjs-jquery (3.2.1-1) ...
Setting up libtexlua52:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Setting up fonts-droid-fallback (1:6.0.1r16-1.1) ...
Setting up libsynctex1:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Setting up libptexenc1:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Setting up tex-common (6.09) ...
update-language: texlive-base not installed and configured, doing nothing!
Setting up poppler-data (0.4.8-2) ...
Setting up tex-gyre (20160520-1) ...
Setting up preview-latex-style (11.91-1ubuntu1) ...
Setting up fonts-texgyre (20160520-1) ...
Setting up fonts-noto-mono (20171026-2) ...
Setting up fonts-lato (2.0-2) ...
Setting up libcupsfilters1:amd64 (1.20.2-0ubuntu3.1) ...
Setting up libcupsimage2:amd64 (2.2.7-1ubuntu2.9) ...
Setting up libjbig2dec0:amd64 (0.13-6) ...
```

```
Setting up ruby-did-you-mean (1.2.0-2) ...
Setting up t1utils (1.41-2) ...
Setting up ruby-net-telnet (0.1.1-2) ...
Setting up libijs-0.35:amd64 (0.35-13) ...
Setting up rubygems-integration (1.11) ...
Setting up libpotrace0 (1.14-2) ...
Setting up javascript-common (11) ...
Setting up ruby-minitest (5.10.3-1) ...
Setting up libzzip-0-13:amd64 (0.13.62-3.1ubuntu0.18.04.1) ...
Setting up libgs9:amd64 (9.26~dfsg+0-0ubuntu0.18.04.17) ...
Setting up libtexluajit2:amd64 (2017.20170613.44572-8ubuntu0.1) ...
Setting up fonts-lmodern (2.004.5-3) ...
Setting up ruby-power-assert (0.3.0-1) ...
Setting up texlive-binaries (2017.20170613.44572-8ubuntu0.1) ...
update-alternatives: using /usr/bin/xdvi-xaw to provide /usr/bin/xdvi.bin (xdvi.bi
n) in auto mode
update-alternatives: using /usr/bin/bibtex.original to provide /usr/bin/bibtex (bi
btex) in auto mode
Setting up texlive-base (2017.20180305-1) ...
mktexlsr: Updating /var/lib/texmf/ls-R-TEXLIVEDIST...
mktexlsr: Updating /var/lib/texmf/ls-R-TEXMFMAIN...
mktexlsr: Updating /var/lib/texmf/ls-R...
mktexlsr: Done.
tl-paper: setting paper size for dvips to a4: /var/lib/texmf/dvips/config/config-p
aper.ps
tl-paper: setting paper size for dvipdfmx to a4: /var/lib/texmf/dvipdfmx/dvipdfmx-
paper.cfg
tl-paper: setting paper size for xdvi to a4: /var/lib/texmf/xdvi/XDvi-paper
tl-paper: setting paper size for pdftex to a4: /var/lib/texmf/tex/generic/config/p
dftexconfig.tex
Setting up texlive-fonts-recommended (2017.20180305-1) ...
Setting up texlive-plain-generic (2017.20180305-2) ...
Setting up texlive-latex-base (2017.20180305-1) ...
Setting up lmodern (2.004.5-3) ...
Setting up texlive-latex-recommended (2017.20180305-1) ...
Setting up texlive-pictures (2017.20180305-1) ...
Setting up tipa (2:1.3-20) ...
Regenerating '/var/lib/texmf/fmtutil.cnf-DEBIAN'... done.
Regenerating '/var/lib/texmf/fmtutil.cnf-TEXLIVEDIST'... done.
update-fmtutil has updated the following file(s):
        /var/lib/texmf/fmtutil.cnf-DEBIAN
        /var/lib/texmf/fmtutil.cnf-TEXLIVEDIST
If you want to activate the changes in the above file(s),
you should run fmtutil-sys or fmtutil.
Setting up texlive (2017.20180305-1) ...
Setting up texlive-latex-extra (2017.20180305-2) ...
Setting up texlive-xetex (2017.20180305-1) ...
Setting up ruby2.5 (2.5.1-1ubuntu1.12) ...
Setting up ruby (1:2.5.1) ...
Setting up ruby-test-unit (3.2.5-1) ...
Setting up rake (12.3.1-1ubuntu0.1) ...
Setting up libruby2.5:amd64 (2.5.1-1ubuntu1.12) ...
Processing triggers for mime-support (3.60ubuntu1) ...
Processing triggers for libc-bin (2.27-3ubuntu1.6) ...
Processing triggers for man-db (2.8.3-2ubuntu0.1) ...
Processing triggers for fontconfig (2.12.6-0ubuntu2) ...
Processing triggers for tex-common (6.09) ...
Running updmap-sys. This may take some time...
```