```
%reload_ext autoreload
%autoreload 2
%matplotlib inline
```

```
from fastai.text import *
```

```
path = untar_data(URLs.IMDB_SAMPLE)
```

Downloading http://files.fast.ai/data/examples/imdb_sample

```
path.ls()
```

[PosixPath('/root/.fastai/data/imdb_sample/texts.csv')]

```
df = pd.read_csv(path/'texts.csv')
```

```
df.head()
```

	label	text	is_valid
0	negative	Un-bleeping-believable! Meg Ryan doesn't even ...	False
1	positive	This is a extremely well-made film. The acting...	False
2	negative	Every once in a long while a movie will come a...	False
3	positive	Name just says it all. I watched this movie wi...	False
4	negative	This movie succeeds at being one of the most u...	False

```
df['text'][2]
```

'Every once in a long while a movie will come along that will be so awful that I feel

```
data_lm = TextDataBunch.from_csv(path, 'texts.csv')
```

```
data_lm.save()
```

```
data = load_data(path)
```

```
data.show_batch()
```

text

xxbos xxmaj raising xxmaj victor xxmaj vargas : a xxmaj review \n \n xxmaj you know , xxmaj raising xxmaj victor xxmaj vargas is like sticking your hands into a big , steaming bowl of xxunk . xxmaj it 's warm and gooey , but you 're not sure if it feels right . xxmaj try as i might , no matter how warm and gooey xxmaj raising xxmaj

xxbos xxup the xxup shop xxup around xxup the xxup corner is one of the sweetest and most feel - good romantic comedies ever made . xxmaj there 's just no getting around that , and it 's hard to actually put one 's feeling for this film into words . xxmaj it 's not one of those films that tries too hard , nor does it come up with

xxbos xxmaj now that xxmaj che(2008) has finished its relatively short xxmaj australian cinema run (extremely limited xxunk screen in xxmaj sydney , after xxunk) , i can xxunk join both xxunk of " xxmaj at xxmaj the xxmaj movies " in taking xxmaj steven xxmaj soderbergh to task . \n \n xxmaj it 's usually satisfying to watch a film director change his style /

```
data.vocab.itos[:10]
```

```
['xxunk',
 'xxpad',
 'xxbos',
 'xxeos',
 'xxfld',
 'xxmaj',
 'xxup',
 'xxrep',
 'xxwrep',
 'the']
```

```
data.train_ds[0][0]
```

```
Text xxbos xxmaj this is a wonderfully written and well acted psychological drama . x

    a great example of how to make a good film on a small budget , without big studios
```

```
data.train_ds[0][0].data[:10]
```

```
array([   2,    5,   21,   16,   13, 1557,  420,   12,   89,  939])
```

```
data = (TextList.from_csv(path, 'texts.csv', cols='text')
            .split_from_df(col=2)
            .label_from_df(cols=0)
            .databunch())
```

```
bs=48
```

```
path = untar_data(URLs.IMDB)
path.ls()
```

```
      [PosixPath('/root/.fastai/data/imdb/README'),
       PosixPath('/root/.fastai/data/imdb/imdb.vocab'),
       PosixPath('/root/.fastai/data/imdb/tmp_lm'),
       PosixPath('/root/.fastai/data/imdb/unsup'),
       PosixPath('/root/.fastai/data/imdb/train'),
       PosixPath('/root/.fastai/data/imdb/tmp_clas'),
       PosixPath('/root/.fastai/data/imdb/test')]
```

```
(path/'train').ls()
```

```
[PosixPath('/root/.fastai/data/imdb/train/neg'),
 PosixPath('/root/.fastai/data/imdb/train/labeledBow.feat'),
 PosixPath('/root/.fastai/data/imdb/train/pos'),
 PosixPath('/root/.fastai/data/imdb/train/unsupBow.feat')]
```

```
data_lm = (TextList.from_folder(path)
            .filter_by_folder(include=['train', 'test', 'unsup'])
            .split_by_rand_pct(0.1)
            .label_for_lm()
            .databunch(bs=bs))
data_lm.save('data_lm.pkl')
```

```
data_lm = load_data(path, 'data_lm.pkl', bs=bs)
```

```
data_lm.show_batch()
```

idx	
0	the night from a xxmaj mariah xxmaj carey nightmare ! ;) \n \n xxmaj so - i watched it , but i watche it in fast forward . xxmaj you get the picture that most of the cast are mean people who are going t bad ways ' as soon as they are introduced . \n \n xxmaj th
1	this movie affected the audience that much . \n \n i consider this to be equal to xxmaj citizen xxm far as quality movies are concerned . xxmaj this is n't light entertainment . xxbos xxmaj in this h revival " of the xxmaj blaxploitation genre , you have the superstars of xxmaj black ' 70s cinema : xxmaj brown (al
2	. xxmaj it 's just your basic sports movie (although there is debate about whether cheering is xxmaj however , the aspect that really brings the movie down is the seriousness that is laid on ext nearly every scene . xxmaj for example , when trying to convince a skeptical newcomer to join xxmaj torrance says with a cl

```
learn = language_model_learner(data_lm, AWD_LSTM, drop_mult=0.3)
```

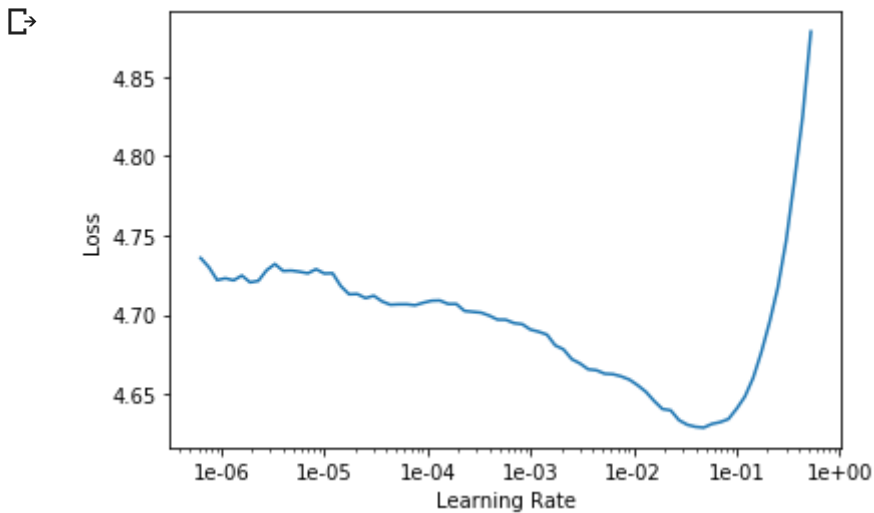Downloading https://s3.amazonaws.com/fast-ai-modelzoo/wt103-fwd

```
learn.lr_find()
```

```
                                            0.00% [0/1 00:00<00:00]
```

epoch	train_loss	valid_loss	accuracy	time

```
                                            1.23% [99/8055 00:30<40:14 11.2372]
LR Finder is complete, type {learner_name}.recorder.plot() to see the graph.
```

```
learn.recorder.plot(skip_end=15)
```



```
learn.fit_one_cycle(1, 1e-2, moms=(0.8,0.7))
```

epoch	train_loss	valid_loss	accuracy	time
0	4.199838	4.021939	0.295500	46:44

```
learn.save('fit_head')
```

```
learn.load('fit_head');
```

```
learn.unfreeze()
```

```
learn.fit_one_cycle(1, 1e-3, moms=(0.8,0.7))
```

epoch	train_loss	valid_loss	accuracy	time
0	3.788058	3.733792	0.329100	52:05

```
learn.save('fine_tuned')
```

```
learn.load('fine_tuned')
```

```
TEXT = "This movie was an absolute waste of time beacause"
N_WORDS = 40
N_SENTENCES = 3
```

```
print("\n".join(learn.predict(TEXT, N_WORDS, temperature=0.75) for _ in range(N_SENTENCES)
```

```
This movie was an absolute waste of time beacause my friends and i watched it just to
This movie was an absolute waste of time beacause by the some actors who are filmmake
This movie was an absolute waste of time beacause . The script was written by David M
```

```
learn.save_encoder('fine_tuned_enc')
```

Classifier for movie reviews

```
path = untar_data(URLs.IMDB)
```

```
data_clas = (TextList.from_folder(path, vocab=data_lm.vocab)
             .split_by_folder(valid='test')
             .label_from_folder(classes=['neg', 'pos'])
             .databunch(bs=bs))
```

```
data_clas.save('data_clas.pkl')
```

⤷

```
data_clas = load_data(path, 'data_clas.pkl', bs=bs)
```

```
data_clas.show_batch()
```

⤷

text
xxbos xxmaj match 1 : xxmaj tag xxmaj team xxmaj table xxmaj match xxmaj bubba xxmaj ray and xxmaj spike xxmaj dudley vs xxmaj eddie xxmaj guerrero and xxmaj chris xxmaj benoit xxmaj bubba xxmaj ray and xxmaj spike xxmaj dudley started things off with a xxmaj tag xxmaj team xxmaj table xxmaj match against xxmaj eddie xxmaj guerrero and xxmaj chris xxmaj benoit . xxmaj according to the rules
xxbos * ! ! - xxup spoilers - ! ! * \n \n xxmaj before i begin this , let me say that i have had both the advantages of seeing this movie on the big screen and of having seen the " xxmaj authorized xxmaj version " of this movie , remade by xxmaj stephen xxmaj king , himself , in 1997 . \n \n xxmaj both
xxbos i thought that xxup rotj was clearly the best out of the three xxmaj star xxmaj wars movies . i find it surprising that xxup rotj is considered the weakest installment in the xxmaj trilogy by many who have voted . xxmaj to me it seemed like xxup rotj was the best because it had the most profound plot , the most suspense , surprises , most xxunk the

```
learn = text_classifier_learner(data_clas, AWD_LSTM, drop_mult=0.5)
learn.load_encoder('fine_tuned_enc')
```

```
learn.lr_find()
```
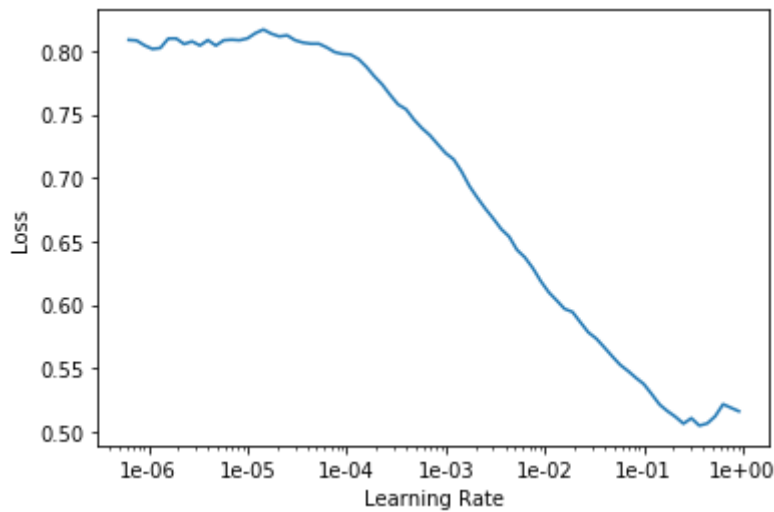
⤷

```
        0.00% [0/1 00:00<00:00]
```

epoch	train_loss	valid_loss	accuracy	time

```
        17.88% [93/520 00:30<02:19 1.7384]
LR Finder is complete, type {learner_name}.recorder.plot() to see the graph.
```

```
learn.recorder.plot()
```

⤷

```
learn.fit_one_cycle(1, 2e-3, moms=(0.8,0.7))
```

epoch	train_loss	valid_loss	accuracy	time
0	0.289168	0.225061	0.908200	04:53

```
learn.save('first')
```

```
learn.load('first')
```

```
learn.freeze_to(-2)
learn.fit_one_cycle(1, slice(1e-2/(2.6**4),1e-2), moms=(0.8,0.7))
```

epoch	train_loss	valid_loss	accuracy	time
0	0.240493	0.182677	0.930840	05:36

```
learn.save('second')
```

```
learn.load('second')
```

```
learn.freeze_to(-3)
learn.fit_one_cycle(1, slice(5e-3/(2.6**4),5e-3), moms=(0.8,0.7))
```

epoch	train_loss	valid_loss	accuracy	time
0	0.203206	0.158264	0.939600	07:26

```
learn.save('third')
```

```
learn.load('third')
```

```
learn.unfreeze()
learn.fit_one_cycle(2, slice(1e-3/(2.6**4),1e-3), moms=(0.8,0.7))
```

epoch	train_loss	valid_loss	accuracy	time
0	0.162605	0.157887	0.940560	09:53
1	0.135893	0.159608	0.941360	08:52

```
learn.predict("My son saw this movie with my husband. He is 13. They are both big movie bu
```

```
(Category pos, tensor(1), tensor([0.0231, 0.9769]))
```

epoch	train_loss	valid_loss	accuracy	time