In [12]:

```
import torchtext
import torch
import torch.nn as nn
from torch.autograd import Variable
import matplotlib.pyplot as plt
import numpy as np
from torchtext.vocab import Vectors
from tqdm.notebook import tqdm

import warnings

def fxn():
    warnings.warn("UserWarning", UserWarning)

with warnings.catch_warnings():
    warnings.simplefilter("ignore")
    fxn()
```

In [13]:

```
# The field that will be used for text data.
text = torchtext.data.Field(include_lengths = True)

# The field that will be used for label data.
label = torchtext.data.Field(sequential=False)

# SST-2 dataset using torchtext.datasets.SST library
train, val, test = torchtext.datasets.SST.splits(text, label, filter_pred=lambda

text.build_vocab(train)

label.build_vocab(train)

train_iter, val_iter, test_iter = torchtext.data.BucketIterator.splits((train, v

# Using FastText word embeddings
url = 'https://dl.fbaipublicfiles.com/fasttext/vectors-wiki/wiki.simple.vec'
text.vocab.load_vectors(vectors=Vectors('wiki.simple.vec', url=url))
```

In [14]:

```
batch = next(iter(train_iter))
```

In [32]:

```
i = 0
for batch in test_iter:
    print(i)
    i = i+1
```

```
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
```

```
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
```

In [24]:

```
batch.text
```

Out[24]:

```
(tensor([[ 3781,    43,    21,   116,    14,  8452,    28,    51,    51,   116],
        [  112,    42,   859,  5124,    82,    28,   717,   240,  8366,     4],
        [ 2091,    26,   212,    68,  1462,  1296,     6,    14,    50,    16],
        [ 5139,   210,    63,   584,   726,   628,   828,  1584, 14591,   498],
        [  569,   160,  1754,    95,     3,     4,     8,  8990,     4,     8],
        [    9, 14749,     8,  2832,     5, 16192,   246,    50,   118,  6344],
        [   23,  2393,    27,     8,  8959,   616,   336,   101,     5,    22],
        [  338,   299,  4137,     7,   756,     2,     3,    36,     4,   117],
        [  685,  3880,    31,  2842,    11,     1,    17,  1106,   994,   698],
        [  833,  2470,   345, 14420,    10,     1,   271,     3,     8,  6384],
        [ 2328,  3875,  1963,     6,    46,     1,     5,  5975,     4,     3],
        [  162,     2,     4,  6167,    11,     1,  1476,     8,  2786,    13],
        [ 3781,     1, 15806,     3,  2762,     1,     5,     7,  1242,  2778],
        [    9,     1,   374, 14973,     2,     1,     4,   767,     5,  3677],
        [    7,     1,     2,     5,     1,     1,   108,   420,  3716,   426],
        [ 3737,     1,     1, 15265,     1,     1,     6,   908,   318,   653],
        [  372,     1,     1,  4260,     1,     1,   212,     5,     3,     4],
        [  353,     1,     1,     3,     1,     1,     2,    22,     4,   158],
        [    2,     1,     1,    13,     1,     1,     1, 15254,    16, 10577],
        [    1,     1,     1,     9,     1,     1,     1,     5,     9,     6],
        [    1,     1,     1,   906,     1,     1,     1,  4288, 10442, 11840],
        [    1,     1,     1,     7,     1,     1,     1,     3,  1073,   218],
        [    1,     1,     1,    66,     1,     1,     1,   195,     5,     2],
        [    1,     1,     1,   231,     1,     1,     1,  4053,  1608,     1],
        [    1,     1,     1,     2,     1,     1,     1,     9,    30,     1],
```

```
          [    1,     1,     1,     1,     1,     1,     1,    36,   123,     1],
          [    1,     1,     1,     1,     1,     1,     1,  1434,  1478,     1],
          [    1,     1,     1,     1,     1,     1,     1,    17,     2,     1],
          [    1,     1,     1,     1,     1,     1,     1,     4,     1,     1],
          [    1,     1,     1,     1,     1,     1,     1,   654,     1,     1],
          [    1,     1,     1,     1,     1,     1,     1,   665,     1,     1],
          [    1,     1,     1,     1,     1,     1,     1,     6,     1,     1],
          [    1,     1,     1,     1,     1,     1,     1,   281,     1,     1],
          [    1,     1,     1,     1,     1,     1,     1,     2,     1,
1]]),
 tensor([19, 12, 15, 25, 14,  8, 18, 34, 28, 23]))
```

In [5]:
```
text.vocab.stoi
```

Out[5]:
```
defaultdict(<bound method Vocab._default_unk_index of <torchtext.vocab.Vocab obj
ect at 0x7f8a68dbe100>>,
            {'<unk>': 0,
             '<pad>': 1,
             '.': 2,
             ',': 3,
             'the': 4,
             'and': 5,
             'of': 6,
             'a': 7,
             'to': 8,
             'is': 9,
             "'s": 10,
             'that': 11,
             'in': 12,
             'it': 13,
             'The': 14,
             'as': 15,
             'film': 16,
             'with': 17,
             'movie': 18,
             'but': 19,
             'for': 20,
             'A': 21,
             'its': 22,
             'an': 23,
             'this': 24,
             'you': 25,
             "n't": 26,
             'be': 27,
             '...': 28,
             'It': 29,
             'on': 30,
             'by': 31,
             '--': 32,
             'has': 33,
             'are': 34,
             'about': 35,
             'more': 36,
             'not': 37,
             'than': 38,
             'at': 39,
             'from': 40,
             'one': 41,
             'have': 42,
             'I': 43,
             'like': 44,
             'his': 45,
             'all': 46,
             "'": 47,
             'so': 48,
```

```
'or': 49,
'-RRB-': 50,
'-LRB-': 51,
'story': 52,
'who': 53,
'into': 54,
'most': 55,
'out': 56,
'does': 57,
'too': 58,
'up': 59,
'just': 60,
'This': 61,
'``': 62,
'comedy': 63,
"''": 64,
'`': 65,
'good': 66,
'will': 67,
'can': 68,
'characters': 69,
'much': 70,
'if': 71,
'even': 72,
'no': 73,
'their': 74,
'funny': 75,
'An': 76,
'some': 77,
'time': 78,
'what': 79,
'way': 80,
'only': 81,
'little': 82,
'your': 83,
'which': 84,
'make': 85,
'work': 86,
'been': 87,
'they': 88,
'would': 89,
'makes': 90,
'never': 91,
'very': 92,
'any': 93,
'he': 94,
'us': 95,
'there': 96,
'enough': 97,
'bad': 98,
'do': 99,
'director': 100,
'was': 101,
'life': 102,
'may': 103,
'movies': 104,
'we': 105,
'through': 106,
'best': 107,
'love': 108,
'drama': 109,
'could': 110,
'There': 111,
':': 112,
'made': 113,
```

```
'something': 114,
'really': 115,
'If': 116,
'own': 117,
'performances': 118,
'well': 119,
'plot': 120,
'films': 121,
'action': 122,
'many': 123,
'should': 124,
'better': 125,
'her': 126,
'when': 127,
'how': 128,
'people': 129,
'without': 130,
'other': 131,
'What': 132,
"'re": 133,
'?': 134,
'look': 135,
'cast': 136,
'also': 137,
'off': 138,
'see': 139,
'humor': 140,
'ever': 141,
'script': 142,
'sense': 143,
'both': 144,
'nothing': 145,
'still': 146,
'two': 147,
'every': 148,
'fun': 149,
'new': 150,
'audience': 151,
'them': 152,
'those': 153,
'character': 154,
'great': 155,
'might': 156,
'feel': 157,
'long': 158,
'kind': 159,
'such': 160,
'-': 161,
'because': 162,
'first': 163,
'performance': 164,
'being': 165,
'get': 166,
'often': 167,
'One': 168,
'entertaining': 169,
"'ve": 170,
'As': 171,
'In': 172,
'here': 173,
'minutes': 174,
'seems': 175,
'tale': 176,
'real': 177,
'But': 178,
```

```
'Hollywood': 179,
'between': 180,
'documentary': 181,
';': 182,
'thriller': 183,
'hard': 184,
'screen': 185,
'few': 186,
'over': 187,
'acting': 188,
'down': 189,
'were': 190,
'heart': 191,
'another': 192,
'feels': 193,
'picture': 194,
'while': 195,
'world': 196,
"'ll": 197,
'almost': 198,
'end': 199,
'less': 200,
'comes': 201,
'quite': 202,
'actors': 203,
'itself': 204,
'take': 205,
'year': 206,
'big': 207,
'come': 208,
'interesting': 209,
'seen': 210,
'these': 211,
'family': 212,
'had': 213,
'romantic': 214,
'things': 215,
'before': 216,
'dialogue': 217,
'material': 218,
'moments': 219,
'rather': 220,
'American': 221,
'You': 222,
'watch': 223,
'Like': 224,
'ca': 225,
'far': 226,
'seem': 227,
'works': 228,
'!': 229,
'find': 230,
'thing': 231,
'after': 232,
'back': 233,
'human': 234,
'me': 235,
'scenes': 236,
'watching': 237,
'years': 238,
'yet': 239,
'While': 240,
'lot': 241,
'making': 242,
'original': 243,
```

```
'ultimately': 244,
'compelling': 245,
'go': 246,
'least': 247,
'right': 248,
'worth': 249,
'cinema': 250,
'fascinating': 251,
'old': 252,
'where': 253,
'man': 254,
'With': 255,
'direction': 256,
'once': 257,
'young': 258,
'gives': 259,
'piece': 260,
'special': 261,
'takes': 262,
'then': 263,
'gets': 264,
'give': 265,
'going': 266,
'keep': 267,
'moving': 268,
'For': 269,
'experience': 270,
'music': 271,
'our': 272,
'subject': 273,
'times': 274,
'anything': 275,
'bit': 276,
'part': 277,
'style': 278,
'think': 279,
'comic': 280,
'history': 281,
'say': 282,
'she': 283,
'Mr.': 284,
'did': 285,
'dull': 286,
'emotional': 287,
'full': 288,
'him': 289,
'laughs': 290,
'same': 291,
'screenplay': 292,
'visual': 293,
'why': 294,
'art': 295,
'Not': 296,
'know': 297,
'point': 298,
'since': 299,
'sometimes': 300,
'though': 301,
'together': 302,
'actually': 303,
'again': 304,
'away': 305,
'entertainment': 306,
'filmmakers': 307,
'flick': 308,
```

```
'genre': 309,
'idea': 310,
'need': 311,
'offers': 312,
'short': 313,
'want': 314,
'Even': 315,
'cinematic': 316,
'dark': 317,
'narrative': 318,
'series': 319,
'show': 320,
'whose': 321,
'All': 322,
'care': 323,
'clever': 324,
'kids': 325,
'whole': 326,
'engaging': 327,
'exercise': 328,
'fans': 329,
'goes': 330,
'manages': 331,
'study': 332,
'title': 333,
'worst': 334,
"'d": 335,
'around': 336,
'charm': 337,
'enjoyable': 338,
'feature': 339,
'premise': 340,
'probably': 341,
'simply': 342,
'women': 343,
'New': 344,
'anyone': 345,
'children': 346,
'done': 347,
'effort': 348,
'filmmaking': 349,
'last': 350,
'matter': 351,
'my': 352,
'place': 353,
'predictable': 354,
'smart': 355,
'And': 356,
'familiar': 357,
'nearly': 358,
'portrait': 359,
'surprisingly': 360,
'sweet': 361,
'three': 362,
'Despite': 363,
'Though': 364,
'always': 365,
'becomes': 366,
'energy': 367,
'feeling': 368,
'horror': 369,
'powerful': 370,
'amusing': 371,
'beautiful': 372,
'pretty': 373,
```

```
'set': 374,
'wo': 375,
'charming': 376,
'effects': 377,
'face': 378,
'looking': 379,
'quirky': 380,
'romance': 381,
'silly': 382,
'strong': 383,
'true': 384,
'trying': 385,
'wit': 386,
"'m": 387,
'French': 388,
'No': 389,
'deeply': 390,
'enjoy': 391,
'fact': 392,
'looks': 393,
'modern': 394,
'plays': 395,
'power': 396,
'rare': 397,
'solid': 398,
'sure': 399,
'tone': 400,
'turns': 401,
'under': 402,
'video': 403,
'beautifully': 404,
'certainly': 405,
'dramatic': 406,
'easy': 407,
'especially': 408,
'half': 409,
'summer': 410,
'believe': 411,
'culture': 412,
'debut': 413,
'intelligent': 414,
'likely': 415,
'put': 416,
'reason': 417,
'recent': 418,
'star': 419,
'theater': 420,
'John': 421,
'everything': 422,
'fine': 423,
'ideas': 424,
'intelligence': 425,
'interest': 426,
'level': 427,
'mess': 428,
'mind': 429,
'now': 430,
'small': 431,
'sort': 432,
'stuff': 433,
'version': 434,
'Its': 435,
'My': 436,
'That': 437,
'above': 438,
```

```
'along': 439,
'camera': 440,
'completely': 441,
'directed': 442,
'each': 443,
'else': 444,
'everyone': 445,
'fresh': 446,
'leave': 447,
'must': 448,
'nor': 449,
'problem': 450,
'proves': 451,
'ride': 452,
'stories': 453,
'suspense': 454,
'truly': 455,
'turn': 456,
'TV': 457,
'adventure': 458,
'already': 459,
'audiences': 460,
'boring': 461,
'ending': 462,
'filmmaker': 463,
'flat': 464,
'high': 465,
'hour': 466,
'lack': 467,
'lacks': 468,
'lives': 469,
'obvious': 470,
'pleasure': 471,
'sad': 472,
'satisfying': 473,
'serious': 474,
'shows': 475,
'spirit': 476,
'storytelling': 477,
'touching': 478,
'Just': 479,
'Michael': 480,
'against': 481,
'attempt': 482,
'beyond': 483,
'either': 484,
'hilarious': 485,
'instead': 486,
'opera': 487,
'perfect': 488,
'play': 489,
'war': 490,
'About': 491,
'Although': 492,
'At': 493,
'Too': 494,
'actor': 495,
'classic': 496,
'complex': 497,
'fails': 498,
'himself': 499,
'hours': 500,
'left': 501,
'light': 502,
'line': 503,
```

```
'melodrama': 504,
'particularly': 505,
'rich': 506,
'satire': 507,
'sequel': 508,
'shot': 509,
'By': 510,
'We': 511,
'become': 512,
'book': 513,
'easily': 514,
'neither': 515,
'past': 516,
'pretentious': 517,
'production': 518,
'social': 519,
'terrific': 520,
'tries': 521,
'viewers': 522,
'ways': 523,
'woman': 524,
'Director': 525,
'animation': 526,
'day': 527,
'different': 528,
'formula': 529,
'head': 530,
'honest': 531,
'images': 532,
'imagination': 533,
'jokes': 534,
'lost': 535,
'message': 536,
'remarkable': 537,
'role': 538,
'scene': 539,
'seeing': 540,
'slow': 541,
'wonderful': 542,
'written': 543,
'To': 544,
'bland': 545,
'cliches': 546,
'coming-of-age': 547,
'engrossing': 548,
'found': 549,
'got': 550,
'impossible': 551,
'inside': 552,
'intriguing': 553,
'leaves': 554,
'mood': 555,
'mystery': 556,
'psychological': 557,
'reality': 558,
'remains': 559,
'running': 560,
'sequences': 561,
'simple': 562,
'themselves': 563,
'thoroughly': 564,
'thoughtful': 565,
'tired': 566,
'writing': 567,
'Love': 568,
```

```
'Time': 569,
'When': 570,
'brilliant': 571,
'comedies': 572,
'crime': 573,
'delivers': 574,
'despite': 575,
'dumb': 576,
'emotionally': 577,
'ends': 578,
'events': 579,
'eyes': 580,
'fairly': 581,
'finally': 582,
'having': 583,
'help': 584,
'hero': 585,
'historical': 586,
'job': 587,
'journey': 588,
'live': 589,
'memorable': 590,
'men': 591,
'passion': 592,
'project': 593,
'soap': 594,
'surprising': 595,
'told': 596,
'viewer': 597,
'violence': 598,
'vision': 599,
'wrong': 600,
'Robert': 601,
'case': 602,
'change': 603,
'cold': 604,
'delightful': 605,
'entirely': 606,
'excellent': 607,
'falls': 608,
'gags': 609,
'given': 610,
'impressive': 611,
'moment': 612,
'otherwise': 613,
'political': 614,
'talent': 615,
'teen': 616,
'tragedy': 617,
'uses': 618,
'usual': 619,
'whether': 620,
'After': 621,
'Disney': 622,
'He': 623,
'Spielberg': 624,
'able': 625,
'adults': 626,
'air': 627,
'captures': 628,
'contrived': 629,
'days': 630,
'death': 631,
'flaws': 632,
'imagine': 633,
```

```
'latest': 634,
'mostly': 635,
'next': 636,
'stupid': 637,
'unsettling': 638,
'2': 639,
'David': 640,
'Has': 641,
'II': 642,
'Nothing': 643,
'acted': 644,
'amount': 645,
'animated': 646,
'appeal': 647,
'appealing': 648,
'attention': 649,
'barely': 650,
'concept': 651,
'deep': 652,
'during': 653,
'entire': 654,
'fantasy': 655,
'gentle': 656,
'getting': 657,
'gone': 658,
'home': 659,
'lead': 660,
'low': 661,
'offer': 662,
'old-fashioned': 663,
'parents': 664,
'period': 665,
'personal': 666,
'possible': 667,
'relationships': 668,
'remake': 669,
'scary': 670,
'sci-fi': 671,
'sentimental': 672,
'sex': 673,
'sit': 674,
'talented': 675,
'tedious': 676,
'tell': 677,
'themes': 678,
'thin': 679,
'try': 680,
'ugly': 681,
'welcome': 682,
'words': 683,
'Allen': 684,
'Big': 685,
'De': 686,
'Full': 687,
'Grant': 688,
'Is': 689,
'More': 690,
'So': 691,
'Some': 692,
'They': 693,
'York': 694,
'across': 695,
'act': 696,
'adaptation': 697,
'ambitious': 698,
```

```
'artist': 699,
'awful': 700,
'bring': 701,
'close': 702,
'creepy': 703,
'depth': 704,
'elements': 705,
'examination': 706,
'felt': 707,
'female': 708,
'final': 709,
'hit': 710,
'laugh': 711,
'magic': 712,
'nice': 713,
'occasionally': 714,
'odd': 715,
'overall': 716,
'plenty': 717,
'quality': 718,
'rarely': 719,
'sharp': 720,
'side': 721,
'someone': 722,
'straight': 723,
'surprise': 724,
'thought': 725,
'understand': 726,
'urban': 727,
'view': 728,
'visually': 729,
'Chan': 730,
'Every': 731,
'Jackson': 732,
'Thing': 733,
'War': 734,
'add': 735,
'approach': 736,
'black': 737,
'casting': 738,
'clear': 739,
'clichés': 740,
'crafted': 741,
'creative': 742,
'difficult': 743,
'emotions': 744,
'ensemble': 745,
'epic': 746,
'execution': 747,
'four': 748,
'frame': 749,
'future': 750,
'game': 751,
'genuine': 752,
'gorgeous': 753,
'issues': 754,
'keeps': 755,
'knows': 756,
'lacking': 757,
'moral': 758,
'pictures': 759,
'pieces': 760,
'promise': 761,
'provocative': 762,
'pure': 763,
```

```
'rest': 764,
'result': 765,
'run': 766,
'single': 767,
'stylish': 768,
'sustain': 769,
'taste': 770,
'touch': 771,
'use': 772,
'utterly': 773,
'watchable': 774,
'weird': 775,
'winning': 776,
'worse': 777,
'2002': 778,
'90': 779,
'America': 780,
'Home': 781,
'Steven': 782,
'apart': 783,
'boy': 784,
'call': 785,
'cheap': 786,
'coming': 787,
'contemporary': 788,
'convincing': 789,
'decent': 790,
'definitely': 791,
'doubt': 792,
'exactly': 793,
'expect': 794,
'eye': 795,
'hand': 796,
'hope': 797,
'important': 798,
'insight': 799,
'liked': 800,
'major': 801,
'masterpiece': 802,
'memory': 803,
'nature': 804,
'needs': 805,
'none': 806,
'novel': 807,
'ones': 808,
'pace': 809,
'pacing': 810,
'perfectly': 811,
'playing': 812,
'previous': 813,
'process': 814,
'puts': 815,
'sensitive': 816,
'sets': 817,
'several': 818,
'situations': 819,
'somewhat': 820,
'start': 821,
'subtle': 822,
'success': 823,
'throughout': 824,
'unexpected': 825,
'wants': 826,
'warm': 827,
'warmth': 828,
```

```
'waste': 829,
'Crush': 830,
'George': 831,
'Moore': 832,
'Movie': 833,
'Ms.': 834,
'Much': 835,
'Murphy': 836,
'Spy': 837,
'ability': 838,
'absolutely': 839,
'ago': 840,
'among': 841,
'begins': 842,
'brings': 843,
'called': 844,
'career': 845,
'conventional': 846,
'cool': 847,
'couple': 848,
'create': 849,
'creates': 850,
'cultural': 851,
'deserves': 852,
'dry': 853,
'earnest': 854,
'episode': 855,
'fare': 856,
'form': 857,
'frequently': 858,
'generic': 859,
'girl': 860,
'hold': 861,
'huge': 862,
'insightful': 863,
'kid': 864,
'leads': 865,
'lots': 866,
'loud': 867,
'middle': 868,
'mix': 869,
'problems': 870,
'quiet': 871,
'saw': 872,
'school': 873,
'sexy': 874,
'slightly': 875,
'sounds': 876,
'strange': 877,
'successful': 878,
'supposed': 879,
'surprises': 880,
'thanks': 881,
'treat': 882,
'tribute': 883,
'truth': 884,
'unfunny': 885,
'viewing': 886,
'witty': 887,
'word': 888,
'working': 889,
'worthy': 890,
'writer': 891,
'writer-director': 892,
'British': 893,
```

```
'Girl': 894,
'Never': 895,
'Oscar': 896,
'Sandler': 897,
'Solondz': 898,
'Watching': 899,
'Williams': 900,
'actress': 901,
'allows': 902,
'battle': 903,
'beauty': 904,
'believable': 905,
'clearly': 906,
'colorful': 907,
'company': 908,
'complete': 909,
'considerable': 910,
'core': 911,
'created': 912,
'depressing': 913,
'derivative': 914,
'drag': 915,
'edge': 916,
'effect': 917,
'endearing': 918,
'equally': 919,
'except': 920,
'exciting': 921,
'figure': 922,
'finds': 923,
'flawed': 924,
'flicks': 925,
'fully': 926,
'funnier': 927,
'guys': 928,
'hardly': 929,
'interested': 930,
'inventive': 931,
'lets': 932,
'living': 933,
'lovely': 934,
'manipulative': 935,
'manner': 936,
'match': 937,
'merely': 938,
'monster': 939,
'name': 940,
'others': 941,
'overcome': 942,
'painful': 943,
'painfully': 944,
'places': 945,
'poetry': 946,
'pointless': 947,
'provides': 948,
'quickly': 949,
'recommend': 950,
'reveals': 951,
'runs': 952,
'second': 953,
'seriously': 954,
'sexual': 955,
'sitting': 956,
'situation': 957,
'slight': 958,
```

```
                 'sophisticated': 959,
                 'spectacle': 960,
                 'stand': 961,
                 'starts': 962,
                 'succeeds': 963,
                 'taking': 964,
                 'target': 965,
                 'terms': 966,
                 'thinking': 967,
                 'toward': 968,
                 'treatment': 969,
                 'twists': 970,
                 'uneven': 971,
                 'unique': 972,
                 'upon': 973,
                 'usually': 974,
                 'vivid': 975,
                 'Both': 976,
                 'Brown': 977,
                 'Bullock': 978,
                 'Does': 979,
                 'From': 980,
                 'How': 981,
                 'Instead': 982,
                 'Kids': 983,
                 'King': 984,
                 'Soderbergh': 985,
                 'Those': 986,
                 'World': 987,
                 'ages': 988,
                 'attempts': 989,
                 'based': 990,
                 'behind': 991,
                 'car': 992,
                 'chemistry': 993,
                 'cinematography': 994,
                 'class': 995,
                 'college': 996,
                 'conclusion': 997,
                 'credit': 998,
                 'cut': 999,
                 ...})
```

In [11]:

```
label.vocab.stoi
```

Out[11]:

```
defaultdict(<bound method Vocab._default_unk_index of <torchtext.vocab.Vocab obj
ect at 0x7f8a58d5d6d0>>,
            {'<unk>': 0, 'positive': 1, 'negative': 2})
```

In [7]:

```
class CBoW(nn.Module):
    def __init__(self, input_size, num_classes, batch_size):
        super(CBoW, self).__init__()
        self.embeddings = nn.Embedding(text.vocab.vectors.size()[0], text.vocab.
        self.embeddings.weight.data.copy_(text.vocab.vectors)
        self.linear = nn.Linear(input_size+1, num_classes, bias = True)

    def forward(self, x):
        x, lengths = x
        lengths = Variable(lengths.view(-1, 1).float())
        embedded = self.embeddings(x)
        average_embed = embedded.mean(0)
        concat = torch.cat([average_embed, lengths], dim = 1) # add lengths as a
        output = self.linear(concat)
        logits = torch.nn.functional.log_softmax(output, dim = 1)
```

```
        return logits

    def predict(self, x):
        logits = self.forward(x)
        return logits.max(1)[1] + 1

    def train(self, train_iter, val_iter, test_iter, num_epochs, learning_rate =
        criterion = torch.nn.NLLLoss()
        optimizer = torch.optim.Adam(model.parameters(), lr = learning_rate)
        loss_vec = []

        for epoch in tqdm(range(1, num_epochs + 1)):
            epoch_loss = 0
            for batch in train_iter:
                x = batch.text
                y = batch.label

                optimizer.zero_grad()

                y_p = self.forward(x)

                loss = criterion(y_p, y-1)
                loss.backward()

                optimizer.step()
                epoch_loss += loss.item()

            self.model = model

            loss_vec.append(epoch_loss / len(train_iter))
            if epoch % 1 == 0:
                acc = self.validate(val_iter)
                print('Epoch {} loss: {} | acc: {}'.format(epoch, loss_vec[epoch
                self.model = model
                self.test(test_iter)

        plt.plot(range(len(loss_vec)), loss_vec)
        plt.xlabel('Epoch')
        plt.ylabel('Loss')
        plt.show()
        print('\nModel trained.\n')
        self.loss_vec = loss_vec
        self.model = model

    def test(self, test_iter):
        "All models should be able to be run with following command."
        upload, trues = [], []
        for batch in test_iter:
            x, y = batch.text, batch.label
            preds = self.predict(x)
            upload += list(preds.data.numpy())
            trues += list(y.data.numpy())

        correct = sum([1 if i == j else 0 for i, j in zip(upload, trues)])
        accuracy = correct / len(trues)
        print('Test Accuracy:', accuracy)

        with open("predictions.txt", "w") as f:
            for u in upload:
                f.write(str(u) + "\n")
```

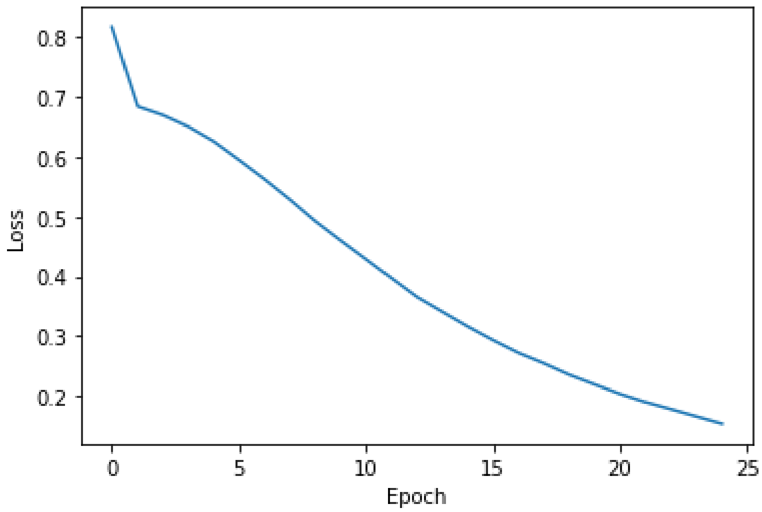
```
    def validate(self, val_iter):
        y_p, y_t, correct = [], [], 0
        for batch in val_iter:
            x, y = batch.text, batch.label
            probs = self.model.predict(x)[:len(y.data.numpy())]
            y_p += list(probs.data.numpy())
            y_t += list(y.data.numpy())
        correct = sum([1 if i == j else 0 for i, j in zip(y_p, y_t)])
        accuracy = correct / len(y_p)
        return accuracy
```

In [8]:

```
model = CBoW(input_size = 300, num_classes = 2, batch_size = 10)

# Hyperparameters tuning experiment 1
model.train(train_iter = train_iter, val_iter = val_iter, test_iter = test_iter,
model.test(test_iter)
```

```
Epoch 1 loss: 0.8159457228948616 | acc: 0.5137614678899083
Test Accuracy: 0.5216913783635365
Epoch 2 loss: 0.6835072180956085 | acc: 0.518348623853211
Test Accuracy: 0.5145524437122461
Epoch 3 loss: 0.6690779727146116 | acc: 0.536697247706422
Test Accuracy: 0.5343218012081274
Epoch 4 loss: 0.6498844925464923 | acc: 0.5768348623853211
Test Accuracy: 0.5678198791872597
Epoch 5 loss: 0.6253119310030358 | acc: 0.6192660550458715
Test Accuracy: 0.6013179571663921
Epoch 6 loss: 0.594331691806027 | acc: 0.6559633027522935
Test Accuracy: 0.641954969796815
Epoch 7 loss: 0.5624791032317057 | acc: 0.694954128440367
Test Accuracy: 0.6963207029104888
Epoch 8 loss: 0.5285249501036081 | acc: 0.7098623853211009
Test Accuracy: 0.7193849533223503
Epoch 9 loss: 0.4924077414449929 | acc: 0.7178899082568807
Test Accuracy: 0.7320153761669412
Epoch 10 loss: 0.460141643795639 | acc: 0.7396788990825688
Test Accuracy: 0.7506864360241625
Epoch 11 loss: 0.4286809805457647 | acc: 0.7580275229357798
Test Accuracy: 0.7748489840746843
Epoch 12 loss: 0.3968484426897041 | acc: 0.7614678899082569
Test Accuracy: 0.7792421746293245
Epoch 13 loss: 0.36555752875229525 | acc: 0.7626146788990825
Test Accuracy: 0.785282811641955
Epoch 14 loss: 0.34080601537417127 | acc: 0.7637614678899083
Test Accuracy: 0.7951674903898956
Epoch 15 loss: 0.31624254827640647 | acc: 0.768348623853211
Test Accuracy: 0.8028555738605162
Epoch 16 loss: 0.29341575132973624 | acc: 0.7694954128440367
Test Accuracy: 0.8072487644151565
Epoch 17 loss: 0.2722940788297467 | acc: 0.7694954128440367
Test Accuracy: 0.8099945085118067
Epoch 18 loss: 0.2549877142684394 | acc: 0.7740825688073395
Test Accuracy: 0.8132894014277869
Epoch 19 loss: 0.23602413653121518 | acc: 0.7798165137614679
Test Accuracy: 0.814936847885777
Epoch 20 loss: 0.219997654437502 | acc: 0.7821100917431193
Test Accuracy: 0.8176825919824272
Epoch 21 loss: 0.20399890673169166 | acc: 0.7821100917431193
Test Accuracy: 0.8176825919824272
Epoch 22 loss: 0.19029189396455798 | acc: 0.783256880733945
Test Accuracy: 0.8193300384404174
Epoch 23 loss: 0.1785978353340071 | acc: 0.7786697247706422
Test Accuracy: 0.8220757825370676
```

```
Epoch 24 loss: 0.16662086854080502 | acc: 0.7798165137614679
Test Accuracy: 0.8231740801757276
Epoch 25 loss: 0.15470783640980462 | acc: 0.7821100917431193
Test Accuracy: 0.8204283360790774
```

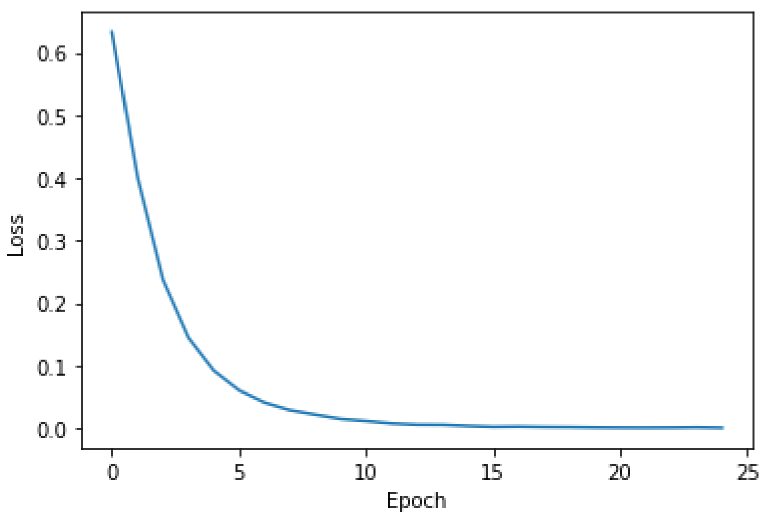



```
Model trained.

Test Accuracy: 0.8204283360790774
```

In [9]:
```
model = CBoW(input_size = 300, num_classes = 2, batch_size = 10)

# Hyperparameters tuning experiment 2
model.train(train_iter = train_iter, val_iter = val_iter, test_iter = test_iter,
model.test(test_iter)
```

```
Epoch 1 loss: 0.633061330959287 | acc: 0.7591743119266054
Test Accuracy: 0.7633168588687534
Epoch 2 loss: 0.4007037608006786 | acc: 0.7717889908256881
Test Accuracy: 0.8099945085118067
Epoch 3 loss: 0.23698939140778402 | acc: 0.7878440366972477
Test Accuracy: 0.8237232289950577
Epoch 4 loss: 0.1443853978285663 | acc: 0.7798165137614679
Test Accuracy: 0.8248215266337178
Epoch 5 loss: 0.09155391052236102 | acc: 0.7717889908256881
Test Accuracy: 0.8160351455244371
Epoch 6 loss: 0.06014263475620971 | acc: 0.7775229357798165
Test Accuracy: 0.8198791872597474
Epoch 7 loss: 0.04058418700937855 | acc: 0.7740825688073395
Test Accuracy: 0.8154859967051071
Epoch 8 loss: 0.028486313310604392 | acc: 0.7775229357798165
Test Accuracy: 0.8023064250411862
Epoch 9 loss: 0.021354635747024592 | acc: 0.7786697247706422
Test Accuracy: 0.7995606809445359
Epoch 10 loss: 0.014439292155615254 | acc: 0.7694954128440367
Test Accuracy: 0.7957166392092258
Epoch 11 loss: 0.011154838949460363 | acc: 0.768348623853211
Test Accuracy: 0.7885777045579352
Epoch 12 loss: 0.007193897348868978 | acc: 0.7672018348623854
Test Accuracy: 0.7869302580999451
Epoch 13 loss: 0.005261975016145706 | acc: 0.7660550458715596
Test Accuracy: 0.7781438769906645
Epoch 14 loss: 0.005041859033662143 5 | acc: 0.768348623853211
Test Accuracy: 0.7764964305326744
Epoch 15 loss: 0.0032049972914436637 | acc: 0.7637614678899083
Test Accuracy: 0.7753981328940143
Epoch 16 loss: 0.0018541449638690248 | acc: 0.7580275229357798
```

```
Test Accuracy: 0.7721032399780341
Epoch 17 loss: 0.002211964448569715 | acc: 0.7626146788990825
Test Accuracy: 0.7655134541460736
Epoch 18 loss: 0.0016042609368406362 | acc: 0.7591743119266054
Test Accuracy: 0.7682591982427238
Epoch 19 loss: 0.0013596250170127282 | acc: 0.7603211009174312
Test Accuracy: 0.7616694124107634
Epoch 20 loss: 0.0008309494718260717 | acc: 0.7637614678899083
Test Accuracy: 0.7677100494233937
Epoch 21 loss: 0.0005102189812849781 | acc: 0.7568807339449541
Test Accuracy: 0.7677100494233937
Epoch 22 loss: 0.00041083408570177126 | acc: 0.7614678899082569
Test Accuracy: 0.7649643053267435
Epoch 23 loss: 0.0005550905099747466 | acc: 0.7557339449541285
Test Accuracy: 0.7644151565074135
Epoch 24 loss: 0.0009647159887313473 | acc: 0.7637614678899083
Test Accuracy: 0.7638660076880834
Epoch 25 loss: 0.00024962132553279 | acc: 0.7660550458715596
Test Accuracy: 0.7649643053267435
```




```
Model trained.

Test Accuracy: 0.7649643053267435
```