sentence_level_genetic_attack_demo

May 11, 2020

1 Demostrate Sentence-level Genetic Attack with Sentence Saliency Analysis

[27]:
```python
from sentence_level_genetic_attack import *
import tensorflow as tf
import stanfordnlp
import time
# uncomment if needed
# stanfordnlp.download('en')
nlp = stanfordnlp.Pipeline()

os.environ['TF_CPP_MIN_LOG_LEVEL'] = '3'  # disable tensorflow loggings.
```

```
Use device: cpu
---
Loading: tokenize
With settings:
{'model_path':
'/Users/weifanjiang/stanfordnlp_resources/en_ewt_models/en_ewt_tokenizer.pt',
'lang': 'en', 'shorthand': 'en_ewt', 'mode': 'predict'}
---
Loading: pos
With settings:
{'model_path':
'/Users/weifanjiang/stanfordnlp_resources/en_ewt_models/en_ewt_tagger.pt',
'pretrain_path':
'/Users/weifanjiang/stanfordnlp_resources/en_ewt_models/en_ewt.pretrain.pt',
'lang': 'en', 'shorthand': 'en_ewt', 'mode': 'predict'}
---
Loading: lemma
With settings:
{'model_path':
'/Users/weifanjiang/stanfordnlp_resources/en_ewt_models/en_ewt_lemmatizer.pt',
'lang': 'en', 'shorthand': 'en_ewt', 'mode': 'predict'}
Building an attentional Seq2Seq model…
Using a Bi-LSTM encoder
Using soft attention for LSTM.
```

```
Finetune all embeddings.
[Running seq2seq lemmatizer with edit classifier]
---
Loading: depparse
With settings:
{'model_path':
'/Users/weifanjiang/stanfordnlp_resources/en_ewt_models/en_ewt_parser.pt',
'pretrain_path':
'/Users/weifanjiang/stanfordnlp_resources/en_ewt_models/en_ewt.pretrain.pt',
'lang': 'en', 'shorthand': 'en_ewt', 'mode': 'predict'}
Done loading processors!
---
```

1.1   load model and dataset to attack

```python
[3]: dataset = "imdb"
     model_name = "pretrained_word_cnn"

     model = word_cnn(dataset)
     model_path = r'./runs/{}/{}.dat'.format(dataset, model_name)
     model.load_weights(model_path)
     print("successfully load model")

     # Data label for imdb dataset:
     # [1 0] is negative review
     # [0 1] is positive review

     train_texts, train_labels, test_texts, test_labels = split_imdb_files()
     x_train, y_train, x_test, y_test = word_process(train_texts, train_labels,␣
      ↪test_texts, test_labels, dataset)
     print('successfully load data')
```

```
Build word_cnn model…
WARNING:tensorflow:From
/Library/Frameworks/Python.framework/Versions/3.7/lib/python3.7/site-
packages/tensorflow_core/python/ops/resource_variable_ops.py:1630: calling
BaseResourceVariable.__init__ (from tensorflow.python.ops.resource_variable_ops)
with constraint is deprecated and will be removed in a future version.
Instructions for updating:
If using Keras pass *_constraint arguments to layers.
WARNING:tensorflow:From
/Library/Frameworks/Python.framework/Versions/3.7/lib/python3.7/site-
packages/tensorflow_core/python/ops/nn_impl.py:183: where (from
tensorflow.python.ops.array_ops) is deprecated and will be removed in a future
version.
Instructions for updating:
Use tf.where in 2.0, which has the same broadcast rule as np.where
```

```
successfully load model
Processing IMDB dataset
successfully load data
```

1.2 demostration

Used test sample 12654 for demostration purposes.

In general, test samples which model does not produce very certain predictions are more likely to have successful attack result. In this case, model predict sample 12654 to be negative with probability $= 0.69$. Our attack is harder to success if this probability is very high (i.e. $>= 0.95$).

Other samples that are easy to success: 12654, 902

```
[4]: test_idx = 12654
     xi_text = test_texts[test_idx]
     yi = test_labels[test_idx]
     print("Clean sample: {}".format(xi_text))
     print("Actual label: {}".format(yi))
     print("Model Prediction: {}".format(predict_str(model, xi_text)))
```

```
Clean sample: This film probably would have been good,if they didn't use CGI
(computer generated imagery)for the werewolf scenes.It made the creatures look
fake and the werewolves looked cartoonish.CGI is great for certain effects like
the dinasours in Jurassic Park or Twister.But when we see a film where the
creature must look completely real,CGI is not the way to go.Look at An American
Werewolf in London.No CGI.Just makeup and a mechanical creature and what you
come up with was more realistic than what was shown in the sequel.This film did
offer a few gags that was fun to watch and the humor in this movie seemed to
have drawn me in but it's nothing more than a film that I thought was O.K.And
that's not good enough.In my opinion,An American Werewolf in Paris doesn't hold
up to the original.
Actual label: [1, 0]
WARNING:tensorflow:From
/Library/Frameworks/Python.framework/Versions/3.7/lib/python3.7/site-
packages/keras/backend/tensorflow_backend.py:422: The name tf.global_variables
is deprecated. Please use tf.compat.v1.global_variables instead.

Model Prediction: [0.68962145 0.30292207]
```

1.2.1 demo computing sentence salience

Let $x = s_1 s_2 \ldots s_n$ be a input consists of n sentences. Let y be x's true label. The sentence saliency for sentence s_k is:

$$S(y|s_k) = P(y|x) - P(x|s_1 s_2 \ldots s_{k-1} s_{k+1} \ldots s_n)$$

3

```
[5]: # break input string to list of sentences
     doc = nlp(xi_text)
     sentences = sentence_list(doc)

     # Compute saliency scores
     raw_saliency = sentence_saliency(model, sentences, yi)

     # Compute normalized saliency scores with softmax
     saliency_scores = softmax(raw_saliency, 10)

     print("Raw sentence saliency: {}".format(raw_saliency))
     print("Softmax sentence saliency: {}".format(saliency_scores))
```

../aten/src/ATen/native/BinaryOps.cpp:81: UserWarning: Integer division of
tensors using div or / is deprecated, and in a future release div will perform
true division as in Python 3. Use true_divide or floor_divide (// in Python)
instead.
../aten/src/ATen/native/TensorAdvancedIndexing.cpp:543: UserWarning:
masked_fill_ received a mask with dtype torch.uint8, this behavior is now
deprecated,please use a mask with dtype torch.bool instead.

Raw sentence saliency: [0.09119469 -0.0176931 -0.2083224 -0.03500342
0.00458783 0.02031082
 -0.01235932 0.05155784 0.00589448 0.00220156]
Softmax sentence saliency: [0.22486347 0.07568769 0.01124949 0.06365719 0.094578
0.11068129
 0.0798343 0.15127885 0.0958219 0.09234782]

1.2.2 demo back-translation for sentence rephrasing.

```
[6]: clean_sentence = "I love to play with little furry cats."
     pivot, final = back_translation(clean_sentence, language = 'zh', require_mid =␣
      ↪True)
     print("Clean sentence: {}".format(clean_sentence))
     print("Pivot translation to Chinese: {}".format(pivot))
     print("Final back translation result: {}".format(final))
```

Clean sentence: I love to play with little furry cats.
Pivot translation to Chinese:
Final back translation result: I like to play with furry kittens.

4

1.2.3 demo genetic attack with verbose output

```
[10]: SEED = 332

      random.seed(SEED)
      np.random.seed(SEED)

      adv_example = genetic(xi_text, yi, model, 20, 5, load_cache(), verbose = True)
      print("adv example: {}".format(adv_example))
      print("clean sample: {}".format(predict_str(model, xi_text)))
      print("adv example pred: {}".format(predict_str(model, adv_example)))
```

../aten/src/ATen/native/TensorAdvancedIndexing.cpp:543: UserWarning:
masked_fill_ received a mask with dtype torch.uint8, this behavior is now
deprecated,please use a mask with dtype torch.bool instead.

```
clean sample's prediction: [0.68962145 0.30292207]
target is to make index 1 > 0.5
generation 1
population 0 pred: [0.7135694 0.2822043], changed idx: {8}
population 1 pred: [0.7091027  0.28510296], changed idx: {0}
population 2 pred: [0.7049275 0.2898213], changed idx: {8}
population 3 pred: [0.87993264 0.1161939 ], changed idx: {7}
population 4 pred: [0.7010012  0.29266572], changed idx: {5}
population 5 pred: [0.65208447 0.34678343], changed idx: {0}
population 6 pred: [0.47788876 0.5174164 ], changed idx: {7}
successful adv. example found!
adv example: This film probably would have been good, if they did n't use CGI(
computer generated imagery) for the werewolf scenes. It made the creatures look
fake and the werewolves looked cartoonish. CGI is great for certain effects like
the dinasours in Jurassic Park or Twister. But when we see a film where the
creature must look completely real, CGI is not the way to go. Look at An
American Werewolf in London. No CGI. Just makeup and a mechanical creature and
what you come up with was more realistic than what was shown in the sequel. This
movie gives some fun to watch and the humor in this movie has taken me to but
nothing more than a movie I thought was good And that's not good enough. In my
opinion, An American Werewolf in Paris does n't hold up to the original.
clean sample: [0.68962145 0.30292207]
adv example pred: [0.47788876 0.5174164 ]
```

1.3 Test more samples

Group 1: model makes non-confident predictions. [3481, 19458, 3612, 24755, 3539]

Code to sample group 1:

```
group1 = list()
while len(group1) < 5:
```

```
        index = random.randint(0, x_test.shape[0])
        confidence = model.predict(x_test[index:index + 1])[0]
        if max(confidence) < 0.7 and index not in group1:
            pred_label = list()
            for conf in confidence:
                pred_label.append(round(conf))
            if pred_label[0] == y_test[index][0] and pred_label[1] == y_test[index][1]:
                group1.append(index)
```

```
[29]: group1 = [3481, 19458, 3612, 24755, 3539]
      for index in group1:
          xi_text = test_texts[index]
          yi = test_labels[index]

          print("------------------------------------")
          print("Sample {}: {}".format(index, xi_text))
          print("pred: {}".format(predict_str(model, xi_text)))

          start = time.time()
          adv = genetic(xi_text, yi, model, 40, 1, load_cache(), verbose = False)
          end = time.time()

          print("Adv: {}".format(adv))
          print("pred: {}".format(predict_str(model, adv)))
          print("Time: {} second".format(end - start))
```

```
------------------------------------
Sample 3481: First of all, i am from munich, where this movie takes place, and
believe it or not, there are guys like Erkan and Stefan, including the silly
dialect! I know their comedy show from the beginning and my main fear was, that
the movie is just an assembling of their already known jokes, but it is not! The
jokes are evolved through the story, and make, in their own special way, sense.
But if you absolutely dislike Erkan und Stefan, hands of this movie. Everyone
else - it's worth viewing!
pred: [0.33844942 0.66082346]

../aten/src/ATen/native/TensorAdvancedIndexing.cpp:543: UserWarning:
masked_fill_ received a mask with dtype torch.uint8, this behavior is now
deprecated,please use a mask with dtype torch.bool instead.

Adv: First of all, i am from munich, where this movie takes place, and believe
it or not, there are guys like Erkan and Stefan, including the silly dialect! I
know their comedy show from the beginning and my main fear was, that the movie
is just an assembling of their already known jokes, but it is not! The jokes are
evolved through the story, and make, in their own special way, sense. But if you
absolutely dislike Erkan und Stefan, hands of this movie. Others-worth a look!
pred: [0.67790467 0.3232927 ]
Time: 59.402191162109375 second
------------------------------------
```

Sample 19458: You know, after the first few Chuck Norris movies, I got so I could tell that a movie was produced by Golan-Globus even if I tuned in in the middle, without ever looking at the credits or the title. What's more I could tell it was Golan-Globus within a minute of screen time. Something about the story structure, the goofy relationships between the characters, the mannered dialog, the wooden acting (spiked with the occasional outright terrible performance), the scene tempos and rhythms that made Albert Pyun look like John McTiernan, the paper-thin plots and not-ready-for-prime-time fight choreography…Golan-Globus has been incredibly consistent over the years in style, subject matter and point-of-view.What can you say, it must work for them, since they've produced literally dozens of movies. You go to one of their productions, and you know exactly what you're getting. And it ain't brain food, folks."Ninja 3" is another piece of hackwork in a long line of products from the G-G sausage factory, and offers the typical limited pleasures to the movie-goers' palate. You've got a Bad Ninja, slicing up cops and criminals and anyone else who gets in their way. You've got a Good Ninja, pledged to stop him. You've got a Westerner thrown into the mix so we Americans can identify with him (or her in this case) and be reassured that "We can still beat those pesky Orientals at their own game." You've got a Love Interest (who is usually also the worst actor/ress in the film) fencing with the Hero. You've got your endless string of assaults, assassinations and lingering shots of men gurgling in agony while an arrow or throwing star sticks unconvincingly out of their eye, neck, or chest. You've got your Beefy White Guy/Bodyguards in Suits calling a Ninja a 'Son of A B*tch' and throwing a roundhouse punch, only to get his *ss handed to him. You've got a Final Confrontation between the Good Guy and The Bad Guy which goes on for 20 minutes and just sort of stops like a RoadRunner cartoon instead of reaching a climax or a resolution.Ninja 3 is a little different, in that the plot revolves around a scrappy female athletic type getting possessed by the Bad Ninja, so she ends up killing a lot of the cops and criminals and Beefy White Bodyguards in Suits while under his spell. But all the other elements are there, as formal in their way as a Kabuki play or a Noh drama.I actually thought Lucinda Dickey was pretty likable in this film. She's nicely muscled and curvy, has great cheekbones and some athletic 'ooomph' to her movements, and you can actually suspend belief enough to accept that her character could do some of the feats she pulls off in the movie. She can almost, but not quite, carry this thing. One extra start for her participation and good energy.Naturally, Sho Kusugi is in here, and he pretty much dominates the last 10-15 minutes of the movie. And just to show you how 3rd-rate and uninspired G-G movies are, the director and editor inter-cut the last climactic fight between Kosugi and the Bad Ninja scene with numerous reaction shots of Dickey and her boyfriend watching the life and death battle with an expression of mild bemusement. I'm serious…for all the emotion and reaction they show to the proceedings, they could be looking at a sea turtle in an aquarium at Marineland. I can only imagine how Dickey must have felt when she saw the finished product – she probably wanted to run the editor through with a katana for real because those reaction shots make her look like a complete idiot. An enjoyable waste of time…but it definitely IS a waste of time. Maybe if you are a Sho Kusugi fan, or even a Linda Dickey fan you'd find it worth your while.

pred: [0.6630755 0.33801666]

../aten/src/ATen/native/TensorAdvancedIndexing.cpp:543: UserWarning: masked_fill_ received a mask with dtype torch.uint8, this behavior is now deprecated,please use a mask with dtype torch.bool instead.

Adv: You know, after the first few Chuck Norris movies, I got so I could tell that a movie was produced by Golan- Globus even if I tuned in in the middle, without ever looking at the credits or the title. What's more I could tell it was Golan- Globus within a minute of screen time. There is something about storytelling, the relationships between characters, storytelling, woodworking (evoked by the most awe-inspiring work), performances and music that made Albert Pyun feel like and John McTiernan, the paper-thin and not-so-ready paperwork for the seasonal … Golan- Globus has been incredibly consistent over the years in style, subject matter and point- of- view. What can you say, it must work for them, since they've produced literally dozens of movies. You go to one of their productions, and you know exactly what you're getting. And it ai n't brain food, folks. " Ninja 3" is another piece of hackwork in a long line of products from the G- G sausage factory, and offers the typical limited pleasures to the movie-goers' palate. You've got a Bad Ninja, slicing up cops and criminals and anyone else who gets in their way. You've got a Good Ninja, pledged to stop him. You've got a Westerner thrown into the mix so we Americans can identify with him(or her in this case) and be reassured that" We can still beat those pesky Orientals at their own game." You've got a Love Interest(who is usually also the worst actor/ ress in the film) fencing with the Hero. You've got your endless string of assaults, assassinations and lingering shots of men gurgling in agony while an arrow or throwing star sticks unconvincingly out of their eye, neck, or chest. You've got your Beefy White Guy/ Bodyguards in Suits calling a Ninja a' Son of A B*tch' and throwing a roundhouse punch, only to get his* ss handed to him. You've got a Final Confrontation between the Good Guy and The Bad Guy which goes on for 20 minutes and just sort of stops like a RoadRunner cartoon instead of reaching a climax or a resolution. Ninja 3 is a little different, in that the plot revolves around a scrappy female athletic type getting possessed by the Bad Ninja, so she ends up killing a lot of the cops and criminals and Beefy White Bodyguards in Suits while under his spell. But all the other elements are there, as formal in their way as a Kabuki play or a Noh drama. I actually thought Lucinda Dickey was pretty likable in this film. She's nicely muscled and curvy, has great cheekbones and some athletic' ooomph' to her movements, and you can actually suspend belief enough to accept that her character could do some of the feats she pulls off in the movie. She can almost, but not quite, carry this thing. One extra start for her participation and good energy. Naturally, Sho Kusugi is in here, and he pretty much dominates the last 10 - 15 minutes of the movie. And just to show you how 3 rd- rate and uninspired G-G movies are, the director and editor inter-cut the last climactic fight between Kosugi and the Bad Ninja scene with numerous reaction shots of Dickey and her boyfriend watching the life and death battle with an expression of mild bemusement. I'm serious… for all the emotion and reaction they show to the proceedings, they could be looking at a sea turtle in an aquarium at Marineland. I can only imagine how Dickey must have felt when she saw the finished product- she

probably wanted to run the editor through with a katana for real because those reaction shots make her look like a complete idiot. An enjoyable waste of time… but it definitely IS a waste of time. Maybe if you are a Sho Kusugi fan, or even a Linda Dickey fan you'd find it worth your while.
pred: [0.35254362 0.6512498]
Time: 341.795291185379 second

Sample 3612: I watched this movie out of a lack of anything else on at the time. Quickly the movie grabbed me with the depth of the characters and subtle plot. I was quite surprised to see that the rating given by my satellite provider did not show any stars.I will not give any spoilers to the plot-line or outcome but most of the characters had at least some redeeming value. Until the very last minute I did not know how it would turn out. Frequently, you can pick-up the formula and predict a story direction. I could not do that with this film.I would recommend this as one of the finer films of this genre.
pred: [0.4832231 0.5159818]

../aten/src/ATen/native/TensorAdvancedIndexing.cpp:543: UserWarning: masked_fill_ received a mask with dtype torch.uint8, this behavior is now deprecated,please use a mask with dtype torch.bool instead.

Adv: I watched this movie out of a lack of anything else on at the time. Quickly the movie grabbed me with the depth of the characters and subtle plot. I was quite surprised to see that the rating given by my satellite provider did not show any stars. I will not give any spoilers to the plot- line or outcome but most of the characters had at least some redeeming value. Until the very last minute I did not know how it would turn out. You can often take the formula and make a prediction about the direction of a story. I could not do that with this film. I would recommend this as one of the finer films of this genre.
pred: [0.5778477 0.42860433]
Time: 79.42394495010376 second

Sample 24755: Lone Star Productions sure churned them out in the 1930's, and "Star Packer" has the feel of one of the more rushed ones. John Wayne is U.S. Marshal John Travers, investigating a crooked hoodlum known only as "The Shadow", responsible for stealing cattle, stage holdups and the like, and giving orders from behind the door of a phony wall safe. Yakima Canutt is Travers' trusty Indian sidekick, appropriately named as… well, "Yak".Early on, we find out that Cattlemens Union head Matt Matlock (George pre-Gabby Hayes) is really The Shadow; the dead giveaway is when he offers to buy out his (supposed) niece Anita's half of the Matlock Ranch, since "this is no place for a girl". As Anita, Verna Hillie doesn't have much to do in the film, although in a comic moment, she gets to use a six shooter to blast the butt of one of the villains in a night time scare raid.There are a few curiosities in the film - for one, Wayne's character alternately rides a white horse and a dark horse in the first half of the film. In what could have been a neat device, a hollowed out tree stump used by a henchman is located right in the middle of the street. And finally, the movie doesn't truly live up to it's name, as Sheriff Travers never wears a badge throughout the film, that is, a star packer without a star.The

horse chases, the runaway stage scenes, the stagecoach off the cliff (another curiosity, the horses conveniently get loose from the stage) are all pretty standard stuff. But John Wayne fans will want to see this one for the charisma he displayed early on in his career. For those more critical, the white kerchiefs worn around the forehead by the good guy posse could only mean that they all had a headache.
pred: [0.52035797 0.47726965]

../aten/src/ATen/native/TensorAdvancedIndexing.cpp:543: UserWarning: masked_fill_ received a mask with dtype torch.uint8, this behavior is now deprecated,please use a mask with dtype torch.bool instead.

Adv: Lone Star Productions sure churned them out in the 1930's, and" Star Packer" has the feel of one of the more rushed ones. John Wayne is U.S. Marshal John Travers, investigating a crooked hoodlum known only as" The Shadow", responsible for stealing cattle, stage holdups and the like, and giving orders from behind the door of a phony wall safe. Yakima Canutt is Travers' trusty Indian sidekick, appropriately named as… well," Yak". Early on, we find out that Cattlemens Union head Matt Matlock(George pre-Gabby Hayes) is really The Shadow; the dead giveaway is when he offers to buy out his(supposed) niece Anita's half of the Matlock Ranch, since" this is no place for a girl". As Anita, Verna Hillie does n't have much to do in the film, although in a comic moment, she gets to use a six shooter to blast the butt of one of the villains in a night time scare raid. There are a few curiosities in the film- for one, Wayne's character alternately rides a white horse and a dark horse in the first half of the film. In what could have been a neat device, a hollowed out tree stump used by a henchman is located right in the middle of the street. And finally, the movie does n't truly live up to it's name, as Sheriff Travers never wears a badge throughout the film, that is, a star packer without a star. The horse chases, the runaway stage scenes, the stagecoach off the cliff(another curiosity, the horses conveniently get loose from the stage) are all pretty standard stuff. But John Wayne fans love to see this cartoon, which was shown early in his career. For those more critical, the white kerchiefs worn around the forehead by the good guy posse could only mean that they all had a headache.
pred: [0.3433545 0.64830285]
Time: 127.89625787734985 second

Sample 3539: This is the only Christopher Guest movie that rivals Spinal Tap and Princess Bride for sheer entertainment value, but somehow never gets near the recognition. The plot surrounds the contestants--dogs--and their owners as they venture into the world of competitive dog…OK, it's about a dog show. The owners truly are characters, as one would have to be to be so attached to their dogs. That's really all there is to it, but that makes it funny enough.You'd never be able to convince me that a mock-u-mentary about dog shows would be funny prior to catching the hilarious scene where Levy and O'hara visit Larry Miller's house on TV…but that's really all it takes to convert any doubters. Spinal Tap was non-stop hilarity, joke after joke whereas Best in Show was had a few more lulls (and by that I mean say 3 minute at MOST where something riotously funny doesn't happen), but the big laughs are even bigger.The casting

in this one is great and even the typically out of place in, uh movies in
general Parker Posey does a fine job. In fact, her tirade directed at Ed Begley
Jr. and a pet store owner over a lost dog toy is probably the funniest running
gag of the film.What's amazing about this movie to me is how the writers somehow
managed to weave a plot, simple as it was, around these great jokes so that it
actually felt like it had direction. I guess there's a freedom in having such a
minimal plot. Everyone's role is pretty well crafted here and the characters are
rarely over-the-top. The realism of how pathetic they seem to the outsider is
what makes it funnier than Mighty Wind or the uneven Guffman. I actually
encounter wierdos like this now and then. If you like Guest's stuff at all, you
should definitely own this one.
pred: [0.38134402 0.62349737]

../aten/src/ATen/native/TensorAdvancedIndexing.cpp:543: UserWarning:
masked_fill_ received a mask with dtype torch.uint8, this behavior is now
deprecated,please use a mask with dtype torch.bool instead.

Adv: This is the only Christopher Guest movie that rivals Spinal Tap and
Princess Bride for sheer entertainment value, but somehow never gets near the
recognition. The plot surrounds the contestants-- dogs-- and their owners as
they venture into the world of competitive dog… OK, it's about a dog show. The
owners truly are characters, as one would have to be to be so attached to their
dogs. That's really all there is to it, but that makes it funny enough. You'd
never be able to convince me that a mock- u-mentary about dog shows would be
funny prior to catching the hilarious scene where Levy and O'hara visit Larry
Miller's house on TV… but that's really all it takes to convert any doubters.
Spinal Tap was non-stop hilarity, joke after jokewhereas Best in Show was had a
few more lulls(and by that I mean say 3 minute at MOST where something
riotously funny does n't happen), but the big laughs are even bigger. This
casting is great, even inappropriate, and the Parker Posei movies do a good job.
In fact, her tirade directed at Ed Begley Jr. and a pet store owner over a lost
dog toy is probably the funniest running gag of the film. What's amazing about
this movie to me is how the writers somehow managed to weave a plot, simple as
it was, around these great jokes so that it actually felt like it had direction.
I guess there's a freedom in having such a minimal plot. Everyone's role is
pretty well crafted here and the characters are rarely over- the- top. The
realism of how pathetic they seem to the outsider is what makes it funnier than
Mighty Wind or the uneven Guffman. I actually encounter wierdos like this now
and then. If you like Guest's stuff at all, you should definitely own this one.
pred: [0.50266093 0.5026942]
Time: 131.5625660419464 second

Group 2: model makes confident predictions. [18669, 7928, 14333, 17319, 12710]

Code to sample group 2:

```
group2 = list()
while len(group2) < 5:
    index = random.randint(0, x_test.shape[0])
    confidence = model.predict(x_test[index:index + 1])[0]
```

```
        if max(confidence) > 0.7 and index not in group2:
            pred_label = list()
            for conf in confidence:
                pred_label.append(round(conf))
            if pred_label[0] == y_test[index][0] and pred_label[1] == y_test[index][1]:
                group2.append(index)
```

```
group2 = [18669, 7928, 14333, 17319, 12710]
for index in group2:
    xi_text = test_texts[index]
    yi = test_labels[index]

    print("-----------------------------------")
    print("Sample {}: {}".format(index, xi_text))
    print("pred: {}".format(predict_str(model, xi_text)))

    start = time.time()
    adv = genetic(xi_text, yi, model, 40, 1, load_cache(), verbose = False)
    end = time.time()

    print("Adv: {}".format(adv))
    print("pred: {}".format(predict_str(model, adv)))
    print("Time: {} second".format(end - start))
```