Sentiment Analysis of Steam Reviews

G12

What have we done?

Comparative Study of Sentiment Analysis of Steam Reviews

-   RNN model and BERT model
-   Compare performance and accuracy of sentiment analysis on game reviews from the platform Steam
-   Both models trained on a subset of 50,000 reviews randomly selected from a larger dataset of Steam reviews
-   Equal number of positive and negative reviews

Why choose this project?

Why Choose to Compare Performance on Steam Reviews?

- Steam reviews are different from normal reviews found on for example IMDB
- Steam reviews often contain humor, sarcasm and community specific language while mostly being short
- Interesting to see if the complexity of BERT better learns the characteristics of the reviews, compared to a simpler model

Currently developed game on Steam

Posted: 8 Apr, 2023 @ 7:03am
Updated: 8 Apr, 2023 @ 10:15am

EARLY ACCESS REVIEW

CRABS OF RAIN 2

Was this review helpful? 👍 Yes 👎 No 😊 Funny 🏅 Award

Posted: 17 Sep, 2023 @ 5:24am

EARLY ACCESS REVIEW

I perfer lobster

Was this review helpful? 👍 Yes 👎 No 😊 Funny 🏅 Award

GaboCrabo
244 products in account
8 reviews

5 reviews

EARLY ACCESS REVIEW

POSTED: 3 MARCH

crab

Was this review helpful?

👍 Yes 👎 No 😊 Funny 🏅 Award

38 people found this review helpful
1 person found this review funny

EARLY ACCESS REVIEW

POSTED: 21 FEBRUARY

I wish crabs were real

Was this review helpful?

👍 Yes 👎 No 😊 Funny 🏅 Award

Posted: 13 Oct, 2023 @ 8:54am

EARLY ACCESS REVIEW

fun

Was this review helpful? 👍 Yes 👎 No 😊 Funny 🏅 Award

Examples of Steam reviews

Timpiish
76 products in account
27 reviews

Recommended
1.4 hrs on record (0.9 hrs at review time)

POSTED: 14 FEBRUARY

The Witcher 3: Wild Hunt is hands down one of the best RPGs I've ever played. I've played the whole game on my other account but now i'm gonna play it on this account also, Seriously, this game is a masterpiece in every sense of the word. From the moment you start playing, you're sucked into this incredibly detailed world filled with monsters, magic, and political intrigue.

The storyline is just...wow. It's so rich and immersive, with twists and turns that keep you on the edge of your seat. And the characters? They feel so real, with their own personalities, motivations, and flaws. You really get invested in their journeys, which makes every decision you make feel weighty and important.

And let's talk about the world itself. It's massive, but it never feels empty or boring. There's always something to do, whether it's hunting down monsters, exploring hidden caves, or just chilling in a tavern playing Gwent. And speaking of Gwent, don't even get me started on how addictive that card game is!

Combat is another highlight. It's challenging, but in a good way. You have to be strategic with your swordplay, magic, and potions to take down tougher enemies, and it feels incredibly satisfying when you pull off a tough fight.

Honestly, I could go on and on about how amazing this game is, but I'll just say this: if you haven't played The Witcher 3 yet, do yourself a favor and pick it up. You won't regret it.

I'm trying to get 10k steam points to customize my profile abit. Help a brotha out:D

Was this review helpful?

👍 Yes   👎 No   😊 Funny   🏅 Award

34 people found this review helpful
1 person found this review funny

🦋 4   👊   🐾   🏅 21

Syksar
71 products in account
3 reviews

Recommended
336.9 hrs on record (324.0 hrs at review time)

POSTED: 9 MARCH

From my personal favorite to least favorite:

Mass Effect

The first game in Legendary Edition is absolutely monumental. Its epic story and breathtaking presentation bring every emotional fluctuation I would hope for in such a heroic space adventure. And the music! There are all sorts of games with great music, but very few if not none of them can reach the quality in this game. The revelation of Citadel is no less an enchanted grandeur than that time Frodo enters Rivendell. Not to mention the final battle and from the wreckage scene, the latter being one of the most powerful moments I have ever experienced in video games.

Sure, there are lots of empty worlds and repetitive level designs, but the diverse scenery is actually enjoyable and brings plenty of Star Trek vibes, that is, if the player is not OCD over every collectible. Fortunately, most of them just give money (you will have much surplus without them) and a little experience (can't hit max level in one playthrough anyway), so you won't be missing anything. Nevertheless, there is hardly a more rewarding moment when all these efforts culminate in the final events. If all comes down to one thing, this game delivers a good adventure, which is just a bit too good.

Mass Effect 3

My first complaint is its inconsistent production quality. How can a game boasts some of the best narrative executions, but at the same time being so ridiculous in audio implementation? And where are the great soundtracks? Even the reused ones from the first game feel out-of-place constantly. Don't even start on details like distant husks in tutorial mission, yikes!

The biggest letdown IMO however, is the lack of credibility in world-building contexts. I can sense how much effort the writers make to convey the desperate and people banding together in overall atmosphere, but suddenly every civilian/soldier/refugee from all species start loving each other and talking honor and ready to sacrifice? It just seems too fairy-tale and ideal. It really should take notes from Halo: Reach, which had very similar story contexts and was released two years earlier than this game, but it portrayed a more authentic scene without

Examples of Steam reviews

Scientific information

Scientific Information Used For the Project

- BERT paper (Devlin et al., 2019)
- Planned on using CBOW paper (Liu, 2020), but decided on RNN as our second model.
- Analysed previous comparative studies on sentiment analysis to find:
  - Feasible models to work with
  - Inspiration on how to compare these (ex. metrics, differing sizes of datasets)

Results

BERT

- 50k Reviews
- 5 Epochs
- Home PC with GPU
- 1h+ Time
- Training Accuracy ~43%
- Validation Accuracy ~38-40%

```
CPU times: user 40min 28s, sys: 26min 43s, total: 1h 7min 11s
Wall time: 1h 6min 26s


782/782 [==============================] - 277s 352ms/step
Accuracy:  0.38864
```

```
Epoch 1/5
4000/4000 [==============================] - 829s 200ms/step - loss: -12.6357 - accuracy: 0.3805 - val_loss: -13.2516 - val_accuracy: 0.4028
Epoch 2/5
4000/4000 [==============================] - 786s 196ms/step - loss: -15.7257 - accuracy: 0.4076 - val_loss: -14.7330 - val_accuracy: 0.3680
Epoch 3/5
4000/4000 [==============================] - 773s 193ms/step - loss: -18.1669 - accuracy: 0.4152 - val_loss: -17.4231 - val_accuracy: 0.4116
Epoch 4/5
4000/4000 [==============================] - 792s 198ms/step - loss: -20.8880 - accuracy: 0.4269 - val_loss: -19.4302 - val_accuracy: 0.4054
Epoch 5/5
4000/4000 [==============================] - 793s 198ms/step - loss: -23.4295 - accuracy: 0.4278 - val_loss: -21.3714 - val_accuracy: 0.4054
```

Recurrent Neural Network - RNN

- 50k Reviews
- Google Collab, ~8 min
- Training Accuracy ~95%
- Validation Accuracy ~83%
- Same or better for 10k, 25k
- More Epochs ->
- Val. Acc Same
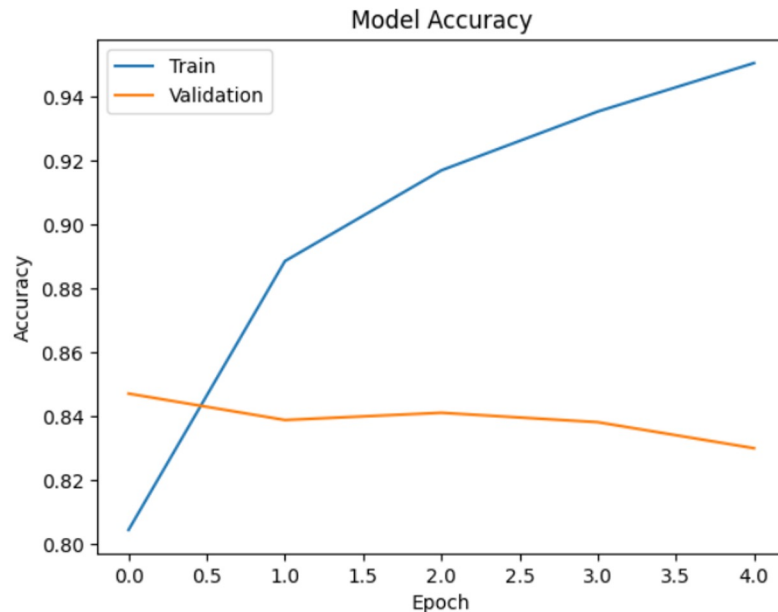
Model Accuracy

```
Epoch 1/5
1104/1104 [==============================] - 108s 95ms/step - loss: 0.4206 - accuracy: 0.8043 - val_loss: 0.3622 - val_accuracy: 0.8470
Epoch 2/5
1104/1104 [==============================] - 99s 90ms/step - loss: 0.2742 - accuracy: 0.8885 - val_loss: 0.3780 - val_accuracy: 0.8388
Epoch 3/5
1104/1104 [==============================] - 100s 90ms/step - loss: 0.2127 - accuracy: 0.9169 - val_loss: 0.3965 - val_accuracy: 0.8410
Epoch 4/5
1104/1104 [==============================] - 99s 90ms/step - loss: 0.1636 - accuracy: 0.9353 - val_loss: 0.4313 - val_accuracy: 0.8381
Epoch 5/5
1104/1104 [==============================] - 95s 86ms/step - loss: 0.1323 - accuracy: 0.9505 - val_loss: 0.4764 - val_accuracy: 0.8299
```

RNN Review Examples

- Real Steam Reviews
- 1.0 -> 100% Positive
- 0.0 -> 100% Negative
- Confidence Variation
- False Positives

```
1/1 [==============================] - 0s 47ms/step
1/1 [==============================] - 0s 73ms/step
Review text: Evolve is really ♥♥♥♥ing fun alone or with friends
Actual score: 1, Predicted score: 1, Confidence: [0.99821544]
--------------------------------------------------
```

```
1/1 [==============================] - 0s 40ms/step
1/1 [==============================] - 0s 45ms/step
Review text: hard for no reason
Actual score: 0, Predicted score: 0, Confidence: [0.01060592]
--------------------------------------------------
```

```
1/1 [==============================] - 0s 47ms/step
1/1 [==============================] - 0s 33ms/step
Review text: Blasphemy. Tropico 3 (with expansion) was awesome, but Tropico 4 is nothing but shameless copycat.
Actual score: 0, Predicted score: 1, Confidence: [0.99673617]
--------------------------------------------------
```

```
1/1 [==============================] - 0s 49ms/step
1/1 [==============================] - 0s 40ms/step
Review text: Look I like cute girls as much as the next otaku but this game is just objectively bad
Actual score: 0, Predicted score: 1, Confidence: [0.6392484]
--------------------------------------------------
```

RNN Testing

- Testing own reviews
- Easily Fooled

```
predict_sentences([
    'Absolutely loved waiting 10 minutes for a single loading screen, really helps me appreciate the art of patience.',
    'Multiplayer mode brings a whole new level of solitude. Never felt more alone while being online.',
    'The developers must be avid historians, considering the game graphics are a tribute to the early 2000s.'
], rounding=False)
```

```
1/1 [==============================] - 0s 31ms/step
array([[0.87765336],
       [0.9882909 ],
       [0.80473894]], dtype=float32)
```

More RNN Testing

- Small input changes ->
- Large Difference in sentiment

```
predict_sentences([
    'Felt like I was in a restaurant, but for my brain. Each level, a delightful new dish.',
    'Felt like I was in a restaurant, but for my brain. Each level, a horrible new dish.'
], rounding=False)
```

```
1/1 [==============================] - 0s 46ms/step
array([[0.9838542 ],
       [0.24267204]], dtype=float32)
```

Conclusions

Our Conclusions From the Project

- BERT model performed worse than expected while RNN model performed very well
- Hardware limitations contributed to the BERT model not being trained on enough data
- RNN models might be better for certain cases when data or hardware is limited