Przedmiot wybieralny III - Przetwarzanie języka naturalnego

Etap 1: Web scrapping recenzji użytkowników (imdb)

Prowadzący:
dr inż. Jakub Osuchowski
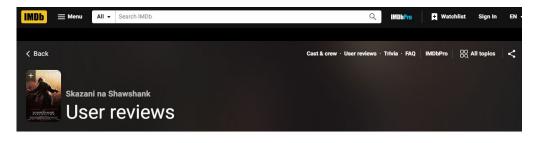
Autorzy:
Paweł Siemiginowski   101450
Paweł Cedzich   101598

TECHNOLOGIE:

‹ Back

Cast & crew · User reviews · Trivia · FAQ   IMDbPro   ⊞ All topics   ↗

Skazani na Shawshank

User reviews

11,459 reviews   ☐ Hide spoilers

Ratings   Show all ▾   Sort by   Featured ▾   ↓   + Review

★ 10/10

Some birds aren't meant to be caged. ›

Spoiler ▾

👍 Helpful · 1.1K   👎 93   ⋮

hitchcockthelegend · Jul 24, 2010 · Permalink

★ 10/10

An incredible movie. One that lives with you. ›

It is no wonder that the film has such a high rating, it is quite literally breathtaking. What can I say that hasn't said before? Not much, it's the story, the acting, the premise, but most of all, this movie is about how it makes you feel. Sometimes you watch a film, and can't remember it days later, this film loves with you, once you've seen it, you don't forget.

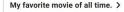The ultimate story of friendship, of hope, and of life, and overcoming adversity.

I understand why so many class this as the best film of all time, it isn't mine, but I get it. If you haven't seen it, or haven't seen it for some time, you need to watch it, it's amazing. 10/10.

More to explore

📷 Photos

Before They Were Famous: Actors' Early Roles

See the gallery

Editorial lists

Related lists created

Tanmay Bhat & The OG
Watchlist
updated 4 months ago ·

Top 100 Movies Bucke

My favorite movie of all time. ›

The Shawshank Redemption has great performances, extremely well written script and story all leading to a deeply emotional climax! One of the best dramas of all time!

👍 Helpful · 127   👎 28   ⋮

0U · Feb 23, 2020 · Permalink

25 more ▾   All ▾

	title	content	rating	helpful_count	not_helpful_count
0	Some birds aren't meant to be caged.	No content	10	1.1K	93
1	An incredible movie. One that lives with you.	It is no wonder that the film has such a high …	10	415	35
2	Don't Rent Shawshank.	I'm trying to save you money; this is the last…	10	1K	120
3	A classic piece of unforgettable film-making.	No content	10	1.3K	169
4	This is How Movies Should Be Made	This movie is not your ordinary Hollywood flic…	10	1K	133
5	Prepare to be moved	I have never seen such an amazing film since I…	No rating	1.6K	246
6	Shawshank Redeems Hollywood	Can Hollywood, usually creating things for ent…	10	1.7K	286
7	the shawshank redemption	The best movie in history and the best ending …	10	164	19
8	Eternal Hope	I've lost count of the number of times I have …	10	221	31
9	The Shawshank Redemption	One of the finest films made in recent years. …	10	928	181
10	IMDb and the Greatest Film of All Time	Two imprisoned men (Tim Robbins and Morgan Fre…	10	110	15
11	Stephen King's best adapted movie	Misery and Stand By Me were the best adaptatio…	10	844	170
12	Mystery	I still can't quite get the enormous reputatio…	7	126	141
13	!BORING!	I do not know, what to say, I am just bored. M…	3	9	13
14	Enthralling, fantastic, intriguing, truly rema…	No content	10	181	33
15	Two movies in one	No content	No rating	1K	240
16	Simply known as "The Movie"	Whenever I talk about this movie with my frien…	10	629	144
17	Simply amazing. The best film of the 90's.	No content	10	604	170
18	All-time prison film classic	No content	10	61	11
19	The absolute GREATEST movie of all time.	No content	10	120	27
20	Simply the best	There's a reason this is rated the #1 . It is …	10	22	2
21	"I'm a convicted murderer who provides sound f…	No content	10	52	10

Przedmiot wybieralny III - Przetwarzanie języka naturalnego

Etap 2: Korekcja błędów

Prowadzący:
dr inż. Jakub Osuchowski

Autorzy:
Paweł Siemiginowski   101450
Paweł Cedzich   101598

TECHNOLOGIE:

pandas

🤗 Transformers

NLTK

Cornell University

arXiv > cs > arXiv:1702.04066

Computer Science > Computation and Language

[Submitted on 14 Feb 2017]

JFLEG: A Fluency Corpus and Benchmark for Grammatical Error Correction

Courtney Napoles, Keisuke Sakaguchi, Joel Tetreault

We present a new parallel corpus, JHU FLuency-Extended GUG corpus (JFLEG) for developing and evaluating grammatical error correction (GEC). Unlike other corpora, it represents a broad range of language proficiency levels and uses holistic fluency edits to not only correct grammatical errors but also make the original text more native sounding. We describe the types of corrections made and benchmark four leading GEC systems on this corpus, identifying specific areas in which they do well and how they can improve. JFLEG fulfills the need for a new gold standard to properly assess the current state of GEC.

Comments: To appear in EACL 2017 (short papers)
Subjects: Computation and Language (cs.CL)
Cite as: arXiv:1702.04066 [cs.CL]
 (or arXiv:1702.04066v1 [cs.CL] for this version)
 https://doi.org/10.48550/arXiv.1702.04066

Submission history
From: Courtney Napoles [view email]
[v1] Tue, 14 Feb 2017 03:47:34 UTC (56 KB)

We gratefully acknowledge support f

HAPPY TRANSFORMER

vennify/t5-base-grammar-correction

```python
def correct_review(review):
    review = fix_spacing(review)
    sentences = sent_tokenize(review)
    corrected_sentences = []

    for sentence in sentences:
        inputs = tokenizer.encode("grammar: " + sentence, return_tensors="pt", max_length=512, truncation=True)
        outputs = model.generate(inputs, max_new_tokens=128)
        corrected_sentence = tokenizer.decode(outputs[0], skip_special_tokens=True)
        corrected_sentences.append(corrected_sentence)

    corrected_review = ' '.join(corrected_sentences)
    return corrected_review
```

content	corrected_content
this is a wierd text. So what can i see, if you see.there it is.	This is a strange text. So what can I see, if you see? There it is.
Hi, my name is Josh.i really... cant fix *****;) it.	Hi, my name is Josh. I really... can't fix it.

Text Compare!

i watch this movie for minimum 5-6 times first to last. so every time i enjoyed as i watching it first time. its a piece of art. and we now know how to entertain with no action or fighting scenes. its a pure drama that anyone can enjoy with family in silence. so i rate this movie 10 out of 10 because of amazing screen play, acting, casting, background scoring editing etc. this is my first movie review so pardon me if i did any mistake.

I watch this movie for minimum 5-6 times first to last. So every time I enjoyed watching it first time. It's a piece of art. And we now know how to entertain with no action or fighting scenes. It's a pure drama that anyone can enjoy with family in silence. So I rate this movie 10 out of 10 because of amazing screenplay, acting, casting, background scoring editing, etc. This is my first movie review, so forgive me if I made any mistakes.

Edit texts ...   Switch texts

Compare!

Clear all

My top film. Hope is good, and it never dies.

dźŚ ,dźŚ ,dźŚ .

My top film. Hope is good, and it never dies.

Przedmiot wybieralny III - Przetwarzanie języka naturalnego

Etap 3: Analiza sentymentu

Prowadzący:
dr inż. Jakub Osuchowski

Autorzy:
Paweł Siemiginowski   101450
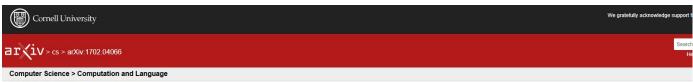Paweł Cedzich   101598

TECHNOLOGIE:

pandas

🤗 Transformers

NLTK

🟣 tabularisai/multilingual-sentiment-analysis 📋

Model Details

- `Model Name:` tabularisai/multilingual-sentiment-analysis
- `Base Model:` distilbert/distilbert-base-multilingual-cased
- `Task:` Text Classification (Sentiment Analysis)
- `Languages:` Supports English plus Chinese (中文), Spanish (Español), Hindi (हिन्दी), Arabic (العربية), Bengali (বাংলা), P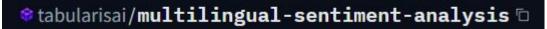ortuguese (Português), Russian (Русский), Japanese (日本語), German (Deutsch), Malay (Bahasa Melayu), Telugu (తెలుగు), Vietnamese (Tiếng Việt), Korean (한국어), French (Français), Turkish (Türkçe), Italian (Italiano), Polish (Polski), Ukrainian (Українська), Tagalog, Dutch (Nederlands), Swiss German (Schweizerdeutsch).
- `Number of Classes:` 5 (*Very Negative, Negative, Neutral, Positive, Very Positive*)

arXiv > cs > arXiv:1910.01108

Computer Science > Computation and Language

[Submitted on 2 Oct 2019 (v1), last revised 1 Mar 2020 (this version, v4)]

DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter

Victor Sanh, Lysandre Debut, Julien Chaumond, Thomas Wolf

As Transfer Learning from large-scale pre-trained models becomes more prevalent in Natural Language Processing (NLP), operating these large models in on-the-edge and/or und In this work, we propose a method to pre-train a smaller general-purpose language representation model, called DistilBERT, which can then be fine-tuned with good performances investigated the use of distillation for building task-specific models, we leverage knowledge distillation during the pre-training phase and show that it is possible to reduce the size of capabilities and being 60% faster. To leverage the inductive biases learned by larger models during pre-training, we introduce a triple loss combining language modeling, distillation pre-train and we demonstrate its capabilities for on-device computations in a proof-of-concept experiment and a comparative on-device study.

Comments: February 2020 - Revision: fix bug in evaluation metrics, updated metrics, argumentation unchanged. 5 pages, 1 figure, 4 tables. Accepted at the 5th Workshop on Energy Efficient Machine Learning an
Subjects: **Computation and Language (cs.CL)**
Cite as: arXiv:1910.01108 **[cs.CL]**
(or arXiv:1910.01108v4 **[cs.CL]** for this version)
https://doi.org/10.48550/arXiv.1910.01108 ⓘ

Submission history
From: Victor Sanh [view email]
[v1] Wed, 2 Oct 2019 17:56:28 UTC (275 KB)
[v2] Wed, 16 Oct 2019 14:52:02 UTC (275 KB)
[v3] Fri, 24 Jan 2020 16:58:52 UTC (276 KB)
[v4] Sun, 1 Mar 2020 02:57:50 UTC (276 KB)

```python
def analyze_sentiment(review):
    if len(review) <= 512:
        sentiments = predict_sentiment([review])
    else:
        chunks = [review[i:i+512] for i in range(0, len(review), 512)]
        sentiments = []
        for chunk in chunks:
            sentiments.extend(predict_sentiment([chunk]))
```

```python
sentiment_scores = [sentiment_values[sentiment] for sentiment in sentiments]
average_sentiment_score = sum(sentiment_scores) / len(sentiment_scores)

# Map average score back to sentiment label
if average_sentiment_score <= -1:
    average_sentiment_label = "Very Negative"
elif average_sentiment_score <= -0.25:
    average_sentiment_label = "Negative"
elif average_sentiment_score < 0.25:
    average_sentiment_label = "Neutral"
elif average_sentiment_score < 1:
    average_sentiment_label = "Positive"
else:
    average_sentiment_label = "Very Positive"
```

Sentence Sentiment

title	content	rating	helpful_co	not_helpf	corrected_	sentiment
IMDB #2 q	This movi	No rating	2	1	This movi	['Negative', 'Very Positive', 'Very Positive', 'Very Positive']
Just Fabul	Morgan Fr	10	1	0	Morgan Fr	['Very Positive', 'Very Positive', 'Positive', 'Positive', 'Very Positive', 'Very Positive', 'Positive', 'Positive', 'Positive', 'Neutral', 'Very Positive', 'Very Positive', 'Very Positive', 'Neutral']
An excelle	When I firs	10	4	2	When I firs	['Neutral', 'Negative', 'Negative', 'Negative', 'Positive', 'Neutral', 'Negative', 'Very Negative', 'Very Negative', 'Positive', 'Positive', 'Neutral', 'Very Positive', 'Neutral', 'Neutral', 'Very Positive', 'Very Positive', 'Neutral', 'Very Positive']
A Film Tha	A powerfu	10	1	0	A powerfu	['Very Positive', 'Very Positive', 'Neutral', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive']
What mov	This is exa	9	1	0	This is exa	['Very Positive']
A fine mov	Good Mov	7	2	2	Good Mov	['Positive', 'Very Positive', 'Neutral', 'Neutral', 'Neutral', 'Neutral', 'Neutral', 'Positive', 'Very Negative', 'Positive', 'Neutral', 'Very Positive', 'Negative']
A Good Mc	In 1947, ba	10	1	0	In 1947, ba	['Very Negative', 'Positive', 'Positive', 'Neutral', 'Neutral', 'Very Negative', 'Very Negative', 'Negative', 'Neutral', 'Positive', 'Neutral', 'Very Negative', 'Neutral', 'Very Negative', 'Neutral', 'Negative', 'Positive', 'Negative', 'Positive', 'Very Negative', 'Neutral'
Best Movic	This is the	10	1	0	This is the	['Very Positive', 'Positive', 'Neutral', 'Neutral', 'Neutral', 'Positive', 'Negative', 'Very Positive', 'Positive', 'Very Positive', 'Very Positive', 'Neutral', 'Positive', 'Negative', 'Very Positive']
Lifted fron	Though th	7	1	1	Though th	['Neutral', 'Very Positive']
Should be	This film i	8	1	0	This film i	['Very Positive', 'Neutral', 'Negative', 'Neutral', 'Neutral', 'Very Negative', 'Very Positive', 'Neutral']
Great	The movie	10	2	1	The movie	['Positive', 'Very Positive', 'Very Negative', 'Positive']
An incredi	It's no wor	9	1	0	It's no wor	['Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Neutral', 'Positive', 'Very Positive', 'Very Positive', 'Very Positive']
A harsh ta	I have see	10	2	1	I have see	['Neutral', 'Neutral', 'Neutral', 'Very Negative', 'Very Negative', 'Positive', 'Neutral', 'Very Positive', 'Positive', 'Very Positive', 'Very Negative', 'Negative', 'Neutral', 'Neutral', 'Very Negative', 'Very Positive', 'Neutral', 'Neutral']
No words	I actually	10	1	0	I actually	['Very Negative', 'Very Positive']
A beautifu	If intentio	10	1	0	If intentio	['Neutral', 'Very Positive', 'Positive']
Great mov	I fully agre	10	1	0	I fully agre	['Positive', 'Positive', 'Very Positive']
Its the nur	This movic	10	1	0	This movic	['Very Positive', 'Positive', 'Very Negative', 'Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Neutral', 'Very Positive', 'Very Positive']
Best movic	This is hov	6	1	1	This is hov	['Very Positive']
My take or	I was ama	10	1	0	I was ama	['Very Positive', 'Negative', 'Neutral', 'Neutral', 'Very Positive', 'Neutral', 'Very Positive', 'Neutral']
Not the be	It is not	7	0	0	It is not m	['Neutral', 'Neutral', 'Neutral', 'Negative', 'Positive', 'Neutral']
Not bad, b	This is not	7	1	1	This is not	['Neutral', 'Very Positive', 'Neutral', 'Very Positive', 'Neutral', 'Positive', 'Neutral', 'Neutral', 'Very Negative', 'Very Negative', 'Positive', 'Neutral', 'Positive', 'Negative', 'Neutral', 'Very Positive', 'Very Negative', 'Neutral', 'Neutral', 'Positive', 'Positive
A good filr	"The Shaw	No rating	1	1	"The Shaw	['Neutral', 'Neutral', 'Very Positive', 'Neutral', 'Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive']
One of the	This move	10	1	0	This move	['Positive']
Must Watc	"The Shaw	10	1	0	"The Shaw	['Very Positive', 'Neutral', 'Very Positive', 'Neutral', 'Very Positive']
Life has nc	This movic	10	2	1	This movic	['Positive', 'Very Positive', 'Very Negative', 'Neutral', 'Very Negative', 'Very Positive', 'Very Positive', 'Neutral']
A bit overh	Decent en	7	14	11	Decent en	['Positive', 'Very Positive', 'Neutral']
Very well b	Strong, me	7	2	1	Strong, me	['Very Positive', 'Very Positive', 'Neutral', 'Negative', 'Negative', 'Negative', 'Neutral', 'Neutral', 'Very Positive', 'Neutral', 'Neutral', 'Neutral']
A Prison S	I thought I	No rating	2	3	I thought I	['Very Negative', 'Neutral', 'Neutral', 'Neutral', 'Neutral', 'Very Negative', 'Positive', 'Neutral', 'Negative', 'Very Positive', 'Neutral', 'Neutral']
Excellent!	The Shaws	10	1	0	The Shaws	['Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive', 'Very Positive']
Good, but	Let me sta	5	3	4	Let me sta	['Positive', 'Very Positive', 'Neutral', 'Positive', 'Neutral', 'Neutral', 'Very Positive', 'Positive']
Truly brilli	I know 'Be	10	5	3	I know 'Be	['Very Negative', 'Negative', 'Positive', 'Very Negative', 'Positive', 'Very Positive', 'Neutral', 'Very Positive', 'Neutral', 'Neutral', 'Very Positive', 'Neutral', 'Very Positive', 'Neutral', 'Very Positive']
Average m	Average ai	6	1	2	Average ai	['Very Negative', 'Positive']
Over Rated	There is nc	7	101	129	There is nc	['Positive', 'Positive', 'Negative', 'Neutral', 'Neutral', 'Neutral']
The Shaws	Stephen K	10	1	0	Stephen K	['Very Positive', 'Positive', 'Positive', 'Positive', 'Neutral', 'Very Positive', 'Neutral', 'Very Positive', 'Positive', 'Positive', 'Very Positive', 'Very Positive', 'Very Positive']
Solid Film	Throughou	10	1	0	Throughou	['Positive']

Review Sentiment

	A	B	C	D	E	F	G
1	title	content	rating	helpful_co	not_helpfu	corrected_content	sentiment
2	IMDB #2 q	This movi	No rating	2	1	This movie was far ahead of its time. It had the best picture, soundtrack, brilliant acting, descriptive characters, m	Very Positive
3	Just Fabul	Morgan Fr	10	1	0	Morgan Freeman, in one of his most iconic roles, gives an outstanding performance as Red. His deep, soulful narr	Very Positive
4	An excelle	When I firs	10	4	2	When I first saw the Shawshank Redemption, I was in a movie theater on a week night at around 11pm. I had gotte	Positive
5	A Film Tha	A powerfu	10	1	0	A powerful and moving story that showcases the human spirit and its resilience in the face of adversity. The Shaws	Very Positive
6	What mov	This is exa	9	1	0	This is exactly the type of movie the new era cinema needs … Magic!	Very Positive
7	A fine mov	Good Mov	7	2	2	Good Movie. The first time I saw it, when it first came out in 1994, I thought that it was fantastic. But now that I have	Positive
8	A Good Mc	In 1947, ba	10	1	0	In 1947, banker Andy Dufresne (Tim Robbins) is convicted of murdering his wife and her lover, based on circumsta	Negative
9	Best Movie	This is the	10	1	0	This is the best movie I've seen. It has a nice context. I understand that the overall rating should be above 9. 5. Or p	Very Positive
10	Lifted from	Though th	7	1	1	Though the movie is good, it's a copy with little modifications. Watch the original, Birdman of Alcatraz.	Neutral
11	Should be	This film i	8	1	0	This film is exceptional. In terms of reality, it's nonsense, but that shouldn't be a barrier. Personally, I don't think it	Very Positive
12	Great	The movie	10	2	1	The movie was a great success. The main character was very positive in all three aspects. In despair also do not g	Positive
13	An incredi	It's no wor	9	1	0	It's no wonder this film has garnered such a high rating - it is truly breathtaking. The story, the acting, and the over	Very Positive
14	A harsh ta	I have see	10	2	1	I have seen this movie a few times. Each time with the same hope and enthusiasm as before. The main story is obv	Positive
15	No words	I actually	10	1	0	I actually don t have any words to describe it. It is just mind blowing, awesome …	Very Positive
16	A beautifu	If intentior	10	1	0	If intention is what counts, A Dream of Freedom is a great film, especially when it focuses on showing the strong fr	Positive
17	Great mov	I fully agre	10	1	0	I fully agree that The Shawshank Redemption is number one on the TOP list. I can watch that movie everyday and it	Very Positive
18	Its the nur	This movie	10	1	0	This movie is amazing from beginning to end. It takes a really good movie for me to not check to see how much lor	Very Positive
19	Best movie	This is hov	6	1	1	This is how movies should be made, this is such a great movie and I will recommend it to everyone.	Very Positive
20	My take or	I was ama	10	1	0	I was amazed, encaptured and enthralled when I saw this film for the first time. I was flicking idlely through the tel	Very Positive
21	Not the be	It is not	7	0	0	It is not my favourite movie ever, but it works for almost everyone. That is saying a lot. It has emotion and it is diffic	Neutral
22	Not bad, b	This is not	7	1	1	This is not a bad film, and I would recommend renting it. But its admirers who rank it as one of the best films of al	Positive
23	A good filr	"The Shaw	No rating	1	1	"The Shawshank Redemption" is not a depressing story, although I may have made it sound that way. There is a lo	Very Positive
24	One of the	This move	10	1	0	This move was released in the early 90s and still gives you a lot of inspirational prospects!	Positive
25	Must Watc	"The Shaw	10	1	0	"The Shawshank Redemption" is a powerful and moving film that tells the story of Andy Dufresne, a man who is wr	Very Positive
26	Life has no	This movie	10	2	1	This movie is one that gives to the most part of being happy in this life. It is true that you might not be lucky enough	Very Positive
27	A bit over	Decent en	7	14	11	Decent entertainment, somewhat memorable. Why this is seen as such an incredible, sensational work of art is bi	Positive
28	Very well k	Strong, me	7	2	1	Strong, meaningful and everlasting message is what makes this movie so powerful. Hope, life and friendship are e	Positive
29	A Prison S	I thought I	No rating	2	3	I thought I had never seen this movie before and then I happened upon it on TV. Then I realized I was required to se	Neutral
30	Excellent!	The Shaws	10	1	0	The Shawshank Redemption is one of those rare films that stays with you long after the credits roll. The story of An	Very Positive
31	Good, but	Let me sta	5	3	4	Let me start by saying that I always enjoy watching this film. It is perhaps one of the ultimate Hollywood feel-good	Very Positive
32	Truly brilli	I know 'Be	10	5	3	I know 'Best Movie Ever' sounds like such a clichÃ©, a phrase used to death. Many people use it on movies which r	Very Positive
33	Average m	Average ai	6	1	2	Average and overrated. Still entertaining and okay to watch with family.	Neutral
34	Over Rated	There is nc	7	101	129	There is no denying that this is a very good movie. The acting is great and it tells a good story. Personally, I find it a	Positive
35	The Shaws	Stephen K	10	1	0	Stephen King's powerful prison drama story is brought to life with Frank Darabont's "The Shawshank Redemption"	Very Positive
36	Solid Film	Throughoi	10	1	0	Throughout watching this movie (The Shawshank Redemption), I found it really enjoyable saying it wasn't my usua	Positive

Kolejny etap:
Analiza sentymentu recenzji użytkowników.

Dziękujemy za uwagę!

Pytania?