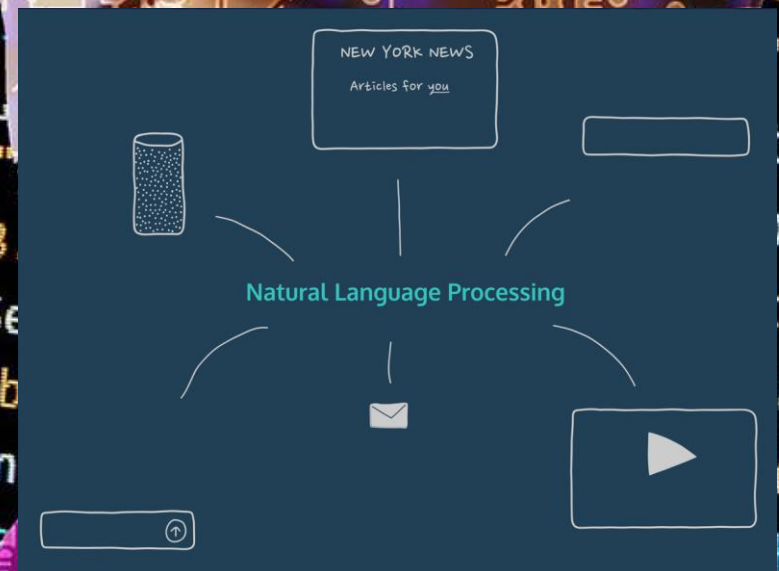
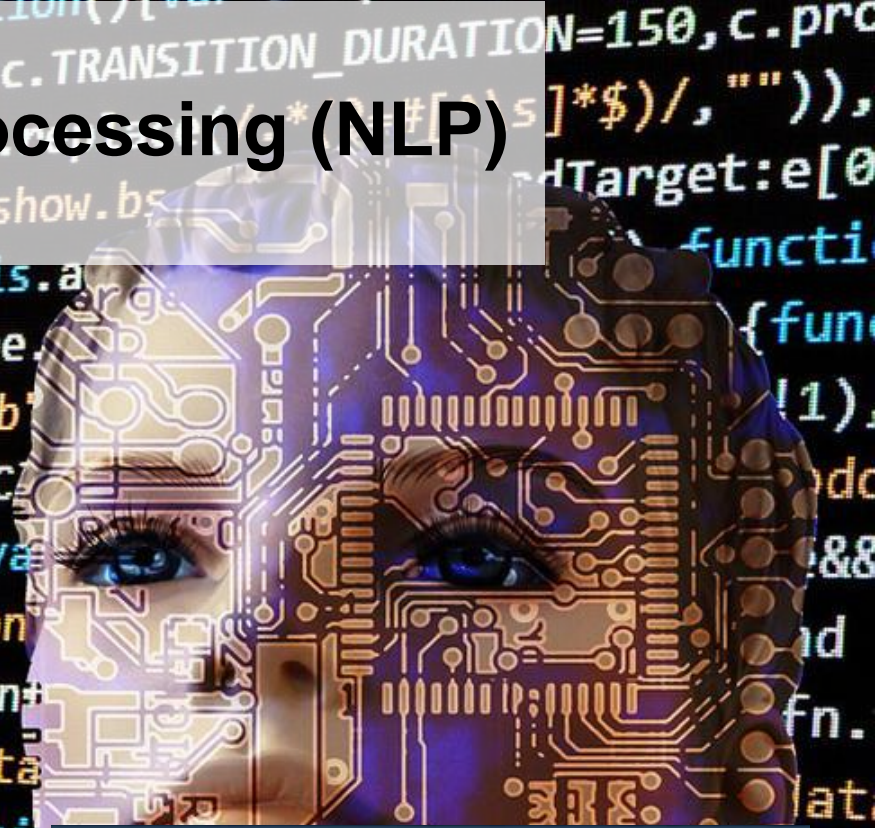


8 Machine Learning II

- ML in Natural Language Processing (NLP)

(1) Motivation & Intro NLP



Which statement on
TLDR is **correct**.

Schwierig- keitsgrad	Art des Wissens	
	Abfragewissen (Vorlesung)	Anwendungswissen (Literatur)
Einfach		
Mittel		
Schwierig		

- a) TLDR is an **Abbreviation** for **Too Long Didn't Read**
- b) TLDR refers to **Texts** with **excessively detailed Information**
- c) **Facebook** is **building** a model called **"TLDR"** to **summarize texts**.
- d) Statements a) and b) are correct.
- e) Statements a), b), and c) are correct.

TLDR: Automatically Summarize in a Click?

tldrthis.com

Special Launch Discount: 30% off your 1st month (any plan) Use Code - LAUNCH30 at checkout

TLDR this

SIGN IN SIGN UP

Automatically summarize any essay in a click.

TLDR This is a web app (+ chrome extension) that can take a painfully long piece of text and sum it up into five key bullet points. Spend less time reading counterproductive stuff and more on doing valuable work.

SUMMARIZE AN ARTICLE

SUMMARY

TLDR NEURAL

For the past decade, we put our faith in technology as the ultimate solver, and any kind of innovation was tied to technological advancement. Silicon Valley has lost some of its halo, and arguably, legitimacy, realizing that the most critical factor in enabling a humane future is not specifically how we relate to one another and the planet we inhabit, but how we relate to ourselves. The most critical factor in enabling a humane future is not specifically how we relate to one another and the planet we inhabit, but how we relate to ourselves.

The dystopia is General AI that is too narrow

But make no mistake: AI is here, and it is going to change everything. These positive changes? And with AI having such a big impact on work, live, play, and even love, are we thinking big enough? How companion in our quest to enable not just our future, but our humanity?

"The business models of the next 10,000 startups are easy to forecast. They will be based on AI. Wired founder Kevin Kelly proclaimed in 2016. That may be true, but at the same time it is disappointing to see that most of the breakthrough applications, from pattern analysis based on massive amounts of reinforcement learning in the style of Deep Mind's Alpha Go to generative adversarial networks performing creative tasks, have been designed to primarily enhance efficiencies (for the enterprise) or convenience (for the consumer).

Before (1700+ words)

Time Saved: 6 mins

1. Digital signatures provide part of the solution, but the main
2. benefits are lost if a trusted third party is still required to prevent double-spending.
3. We propose a solution to the double-spending problem using a peer-to-peer network.
4. The network timestamps transactions by hashing them into an ongoing chain of
5. hash-based proof-of-work, forming a record that cannot be changed without redoing

Facebook is reportedly building a tool called 'TLDR' that will read and summarize news articles for you

Isobel Asher Hamilton Dec 16, 2020, 1:50 PM



Facebook's CEO Mark Zuckerberg Charles Platiau/Reuters

Data to
assembly lines,
QA, shipping products,
tracking orders and
stocking shelves
because fulfilling
the world's every
need is **Everything.**

Show Me

Cambridge Dictionary: TLDR (Too Long; Didn't Read)

Bedeutung von **TLDR** im Englisch



TLDR

(also **TL;DR**, **tldr**, **tl;dr**)



abbreviation for **too long; didn't read**: used to comment on something that someone has written, for example on social media:

- *If a commenter responds to a post with "tl;dr", it expresses an expectation to be entertained without needing to pay attention or to think.*
- *I haven't been able to keep up with this thread (TLDR).*

– Mehr Beispiele

- *There's sometimes a fine line between including every piece of information about the problem, and making it so long that you get TLDR as a response.*
- *TL;DR is a sad symptom of the malfunctions caused by the internet tsunami.*

+ **SMART Vocabulary**: related words and phrases

TLDR: Extreme Summarization of Scientific Documents?

TLDR: Extreme Summarization of Scientific Documents

Isabel Cachola[†]

Kyle Lo[†]

Arman Cohan[†]

Daniel S. Weld^{†‡}

[†]Allen Institute for AI

[‡]Paul G. Allen School of Computer Science & Engineering, University of Washington

{isabelc,kylel,armanc,danw}@allenai.org

We introduce
of extreme s
pers. TLDR
compression
knowledge
domain-spec
on this task,
multi-target
papers. SciTL
and expert-d
are collected
tocol that p
while minim
pose CATTS
ing strategy
exploits titles
CATTS impr
der both auto
tions. Data and code are publicly available at
<https://github.com/allenai/scitldr>.

Dataset	Number of documents	Avg. words in document	Avg. words in summary	Compression ratio	% novel words	Multi-target
<i>Non-scientific documents</i>						
DUC (Over, 2003)	624	441	11	40.1	30.0	yes
NYTimes (Sandhaus, 2008)	655K	549	40	13.7	20.1	no
DailyMail (Hermann et al., 2015)	220K	653	55	11.9	17.0	no
CNN (Hermann et al., 2015)	93K	760	46	16.5	16.8	no
XSUM (Narayan et al., 2018)	226K	431	23	18.7	35.8	no
Newsroom (Grusky et al.)	1.32M	659	27	24.4	26.0	no
BigPatent (Sharma et al., 2019)	1.34M	3.6K	117	30.5	13.6	no
<i>Scientific documents</i>						
CLPubSum (Collins et al., 2017)	10.3K	8.2K	226	36.5	7.7	no
PubMed (Cohan et al., 2018)	133K	3K	203	14.9	10.5	no
ArXiv (Cohan et al., 2018)	215K	4.9K	220	22.5	8.3	no
SciSummNet [†] (Yasunaga et al., 2019)	1.0K	4.7K	150	31.2	7.4	no
TalkSumm [‡] (Lev et al., 2019)	1.7K	4.8K	965	5.0	16.5	no
SciTLDR (ours)	3.2K	5K	21	238.1	15.2	yes

Table 1: Comparison of SciTLDR to existing summarization datasets. (i) SciTLDR provides multiple summary targets unlike other recent summarization datasets. (ii) SciTLDR requires both extreme compression and abstraction, as evidenced by the compression ratio and novelty (% of summary words not in the source document), especially when compared with other scientific summarization datasets.

generation seeks to produce an extreme (single sen

Cachola et.al., 2020: “TLDR: Extreme Summarization of Scientific Documents”

arxiv.org/pdf/2004.15011.pdf