

ITUEvents

Slidin' Videos Challenge



AI for Good
Machine Learning
in 5G Challenge

*Applying machine learning
in communication networks*

Wednesday 27 July 2022

aiforgood.itu.int



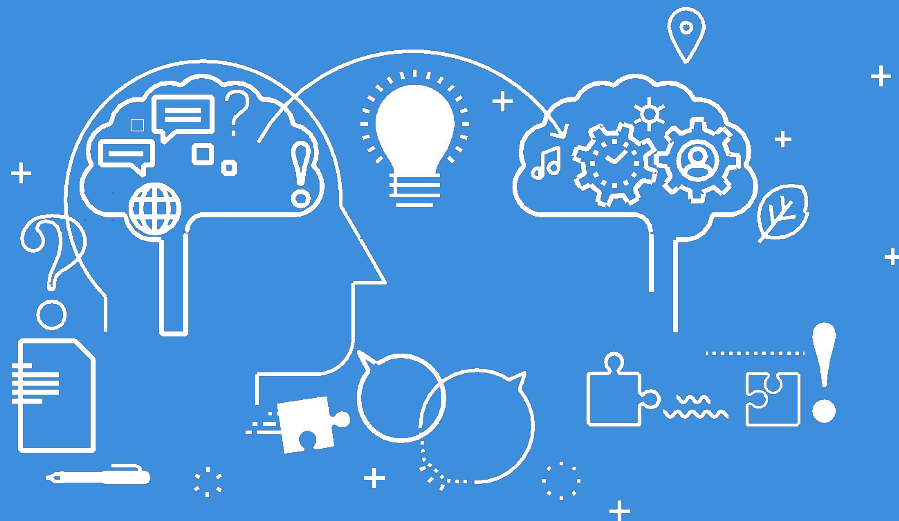


1. Video Chapters

2. Challenge Design

3. Training Dataset

4. Getting started





Video Chapters make structure of video explicit...

Trustworthy AI: Towards Auditable AI Systems | AI FOR GOOD DISCOVERY

The Development of an AI System is an Iterative and Complex Process Which may be Devided Into Phases

A

B

Planning	Data	Training	Evaluation	Operation
Problem statement Available solutions Personnel Data set AI model Learning method Feature selection Run-time system	Acquisition Pre-processing Augmentation Labelling Quality assurance	Parameters Data set partitioning Repetitions Validation Stop criteria Pruning	Setting Relevant input classes White-vs-black-box Performance Repeatability Test cycles	Implementation Hardware embedding Sensors & actuators User & environment interactions

Start of Q&A 40:23

zoom

Federal Office for Information Security

- Helps viewers create a "mind map"
- Encourages creators to structure presentations well

Beyond YouTube:

- Vimeo
- Cincopa
- Others ...

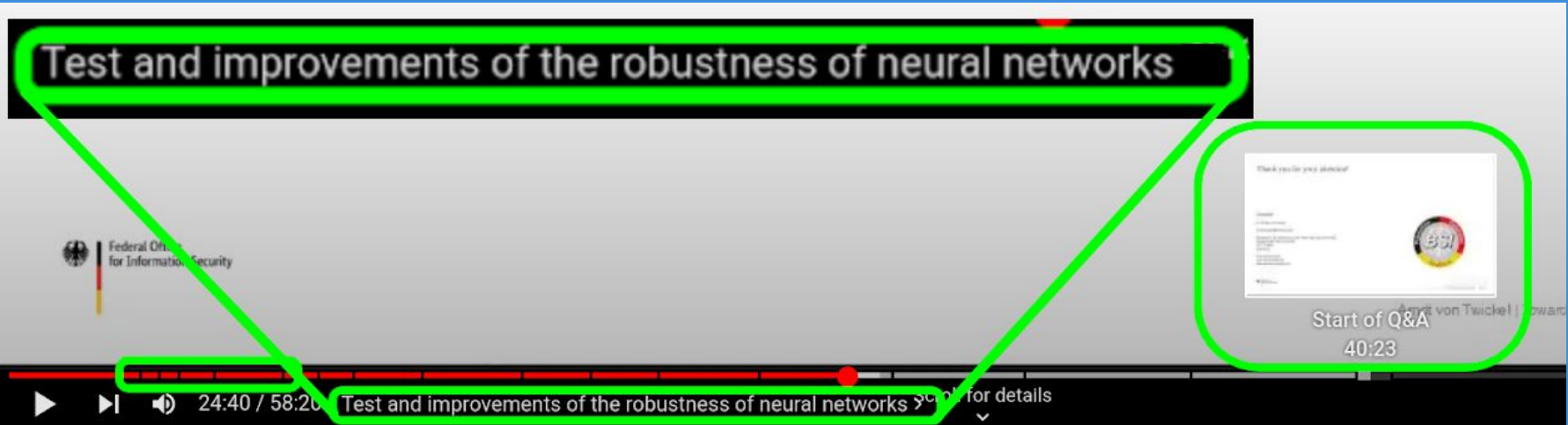


... and provide navigation aid...

- Faster maneuvering
- Assist persons with disabilities, such as visual or hearing impairments

User Interface elements:

1. Chapter marks in the progress bar
2. Title of the current chapter
3. Title of the chapter being previewed (by hover over progress bar)





... and help search within the video

ai for good working together with ai to make new mathematical discoveries X



All News Images Videos Shopping More Tools

About 207,000,000 results (0.49 seconds)

[https://www.youtube.com > watch](https://www.youtube.com/watch)

Alex Davies, DeepMind | WEBINAR - YouTube

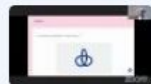


Working together with AI to make new Mathematical Discoveries

| Alex Davies, DeepMind | WEBINAR. Watch later. Share. Copy...

YouTube · AI for Good · 1 month ago

10 key moments in this video



From 13:32
The Study of
Knots



From 15:24
Algebraic
Structure of
Knots



From 16:01
Volume
Conjecture



From 22:29
Representatio
n Theory



Sim
Refl

Metadata is indexable

Search engines deeplink to sections within videos



Limitations of YouTube's auto-generated chapters

- Proprietary
- Not open source - auditable
- Not always available

Test and improvements of the robustness of neural ne...
23:47

0:55 / 58:20 • Intro: Trustworthy AI: Towards Auditable AI Systems >

cleaned by Adblock for Youtube™ [Share](#)

Trustworthy AI: Towards Auditable AI Systems | AI FOR GOOD DISCOVERY

289 views... [Like](#) 4 [Dislike](#) [Share](#) [Download](#) [Clip](#) [Save](#) ...

AI for Good
4.18K subscribers

[ANALYTICS](#) [EDIT VIDEO](#)

Artificial Intelligence (AI) systems such as deep neural networks play a growing role as part of decision and control systems in a plethora of applications. While some of these applications are safety- and security-critical, the use of AI systems pose new challenges with regard to aspects such as IT security, safety, robustness and trustworthiness. To meet these challenges, a generally agreed-upon framework for auditing AI systems throughout their life cycle, comprising evaluation strategies, tools and standards, is required. This is under development but, as of now, only partly ready for practical use. In this talk, focusing on the applications domains autonomous driving and biometrics, the current status of AI system auditing is presented together with open questions and future directions.

SHOWNOTES

- 00:00 Intro: Trustworthy AI: Towards Auditable AI Systems
- 04:00 Responsibilities of BSI in the context of AI
- 04:30 What is the role of BSI in the context of AI?
- 05:00 AI systems are connected and embedded in safety and security-critical applications
- 06:00 CAI differs qualitatively from cIT and sAI

Chapters can be added manually easily

- 1 line per timestamp with title

Shows up automatically as:

1. Chapter Marks
2. Titles



1. Video Chapters

2. Challenge Design

3. Training Dataset

4. Getting started








Key concepts - definitions

- Transitions are generally clearly distinguishable
 - Slide numbers
 - Title changes
 - Presenter hints (e.g. "next slide please")
- Titles are usually obvious, but complex situations exist
 - No title
 - Multiple "titles"
 - Partially covered / off screen
 - Other cases



Example of an obvious slide title

Example knots and invariants

z: Knot	X(z): Geometric Invariants				Y(z): Algebraic Invariants		
	Volume	Chern-Simons	Meridional Translation	...	Signature	Jones Polynomial	...
	2.0299	0	i	...	0	$t^{-2} - t^{-1} + 1 - t + t^2$...
	2.8281	-0.1532	$.7381 + 0.8831i$...	-2	$t - t^2 + 2t^3 - t^4 + t^5 - t^6$...
	3.1640	0.1560	$-.7237 + 1.0160i$...	0	$t^{-2} - t^{-1} + 2 - 2t + t^2 - t^3 + t^4$...

Train an MLP to predict the signature (Y) from the geometric invariants (X)

Source: DeepMind's talk "Working together with AI to make new mathematical discoveries", AI for Good, 27 January 2022



Outliers: herein lies the intelligence

A small video feed of a man wearing headphones, identified as Alex Davies.

Alex Davies

The cover of the November 2023 issue of the journal Nature. The title 'nature' is at the top. The main feature is 'AI-GUIDED INTUITION' with a subtitle 'Machine learning helps inspire mathematicians to derive new conjectures'. The cover art depicts a complex network of lines and nodes, resembling a circuit or a mathematical graph.

"This paper marks the beginning of a new phase in the use of computers in mathematical research."

"I cannot imagine such a mathematician not using these methods, where available, given that not using them means progressing much more slowly and being unable to analyze very complicated objects and very large datasets."

"Such a new approach (even conjectural) to the combinatorial invariance conjecture would certainly be publishable in a top combinatorics journal."

"I judge the separate mathematical paper 'The signature and cusp geometry of hyperbolic knots' likely to be accepted by the top speciality journal in the field (Geometry and Topology) and be a shoe in at the second best one (J. Topology)."

The logo for the journal Nature, a stylized blue 'N' inside a circle.

"AI-Guided Intuition" is a better title than "Nature"



Criteria for evaluation

For each video in the test set, the solution must annotate every slide that was visible (even for a brief time) by specifying

- frame-level timeline boundaries (starting and ending frame numbers)
- (apparent) title of the slide



Submission format

For each video processed by a solution a separate output file with predictions should be generated. The first line in the file must be the header, i.e.:

<code>frame_start,</code>	<code>frame_end,</code>	<code>is_slide,</code>	<code>title</code>
<code>1,</code>	<code>3058,</code>	<code>1,</code>	<code>"AI-based algorithms"</code>
<code>3059,</code>	<code>4550,</code>	<code>1,</code>	<code>"Get involved"</code>
<code>4551,</code>	<code>4980,</code>	<code>1,</code>	<code>"NO_TITLE"</code>

<code>frame_start</code>	initial frame of when slide is visible
<code>frame_end</code>	last video frame containing the slide
<code>is_slide</code>	1 if it is a slide, 0 if something else
<code>title</code>	(apparent) slide title when Is_slide=1

Start frame number must be where the previous chapter ends



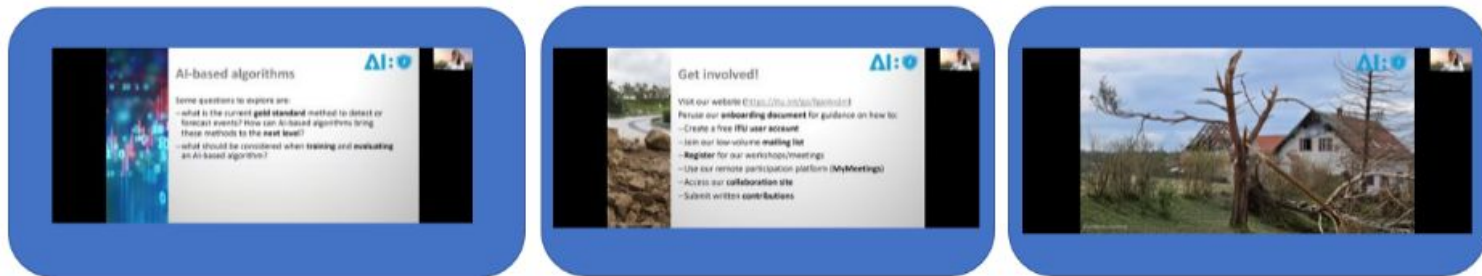
Prediction process

- Slide content may change, not always new slide
- Presenters can switch to non-presentation modes
- **Is_slide = 0** if fragment is not presentation
- Non-slide content can be identified through tracking pixels refresh ratio
- Evaluation metric will focus on accuracy of content and slide transitions only



Prediction process illustration

Frames extracted from a screenshare at 25 FPS



f1

f2 f3

f4 f5

f6

Ground truth labels:

f1, f2, "AI-based algorithms"

f3, f4, "Get involved!"

f5, f6, "NO_TITLE"

where f1, f2, f3, f4, f5, f6 – are frame numbers,
f2 is adjacent to f3, f4 is adjacent to f5



1. Video Chapters
2. Challenge Design
- 3. Training Dataset**
4. Getting started





Training Dataset

- Number of videos: 140
- Total length: ~40 hours
- Number of slide transitions: 2500
- Average length of chapter title ~50 characters

Video files cover the presentation from when speaker started screenshare until they stop it.

Video files vary in

- duration - from several minutes to several hours
- resolution - from 1600x1200 to 3840x2160



Ground truth format

#	frame_start	frame_start	frame_end	frame_end	title1	title2	title3	title4	bonus_title1	bonus_title2
...
11	576	590	1101	101	MVP AI				MVP AI Contribution points	
...								



Floating boundary

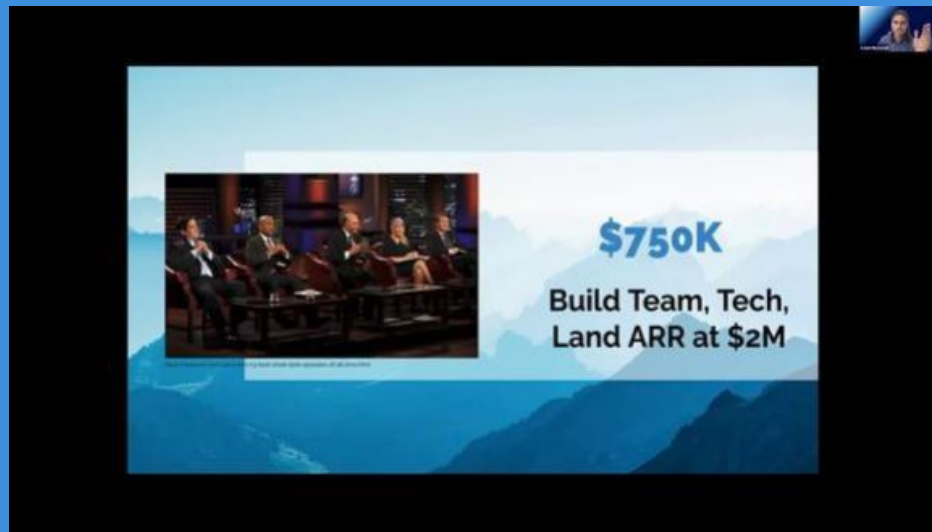


...	starting_frame_begin	starting_frame_end	...
...	397	402	...

Any frame within this range (397 to 426) may be considered as the beginning of a slide



Slide TitleSpecial characters

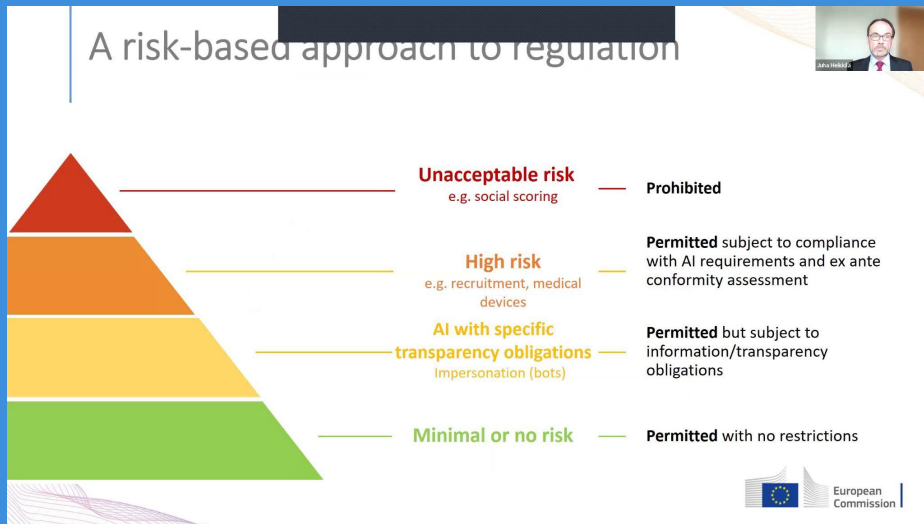


Ground truth			
...	title1	title2	...
	
	S750K	Build Team, Tech, Land, ARR at S2M	
	

Currency signs, greek letters and other special characters will be handled by our evaluation metric



Slide Title Bonus points



Ground truth			
...	title1	...	bonus_title1
	
	A risk-base		A risk-based approach to regulation
	

Use language model and your imagination to earn additional points by fixing “flawed” titles

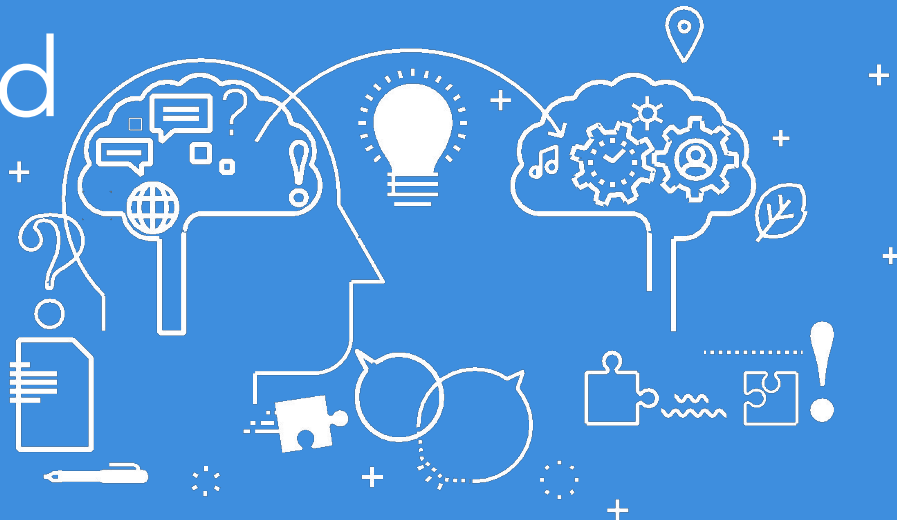


Evaluation workflow

1. Predicted and ground truth slides pairing
2. Predicted boundary check
3. Calculating boundary accuracy
4. Predicted title normalization
5. Predicted title check
6. Calculating title accuracy
7. Weighted average of boundary and title accuracies



1. Video Chapters
2. Challenge Design
3. Training Dataset
4. Getting started





Getting Started

- Training dataset is live
- Submission form is live

<https://challenge.aiforgood.itu.int/match/matchitem/74>



<https://challenge.aiforgood.itu.int/match/matchitem/74>

Any questions? Please contact us / our team at
bastiaan.quast@itu.int / kirill.ekshembeev@itu.int