

Knowledge graphs in neurosymbolic learning approaches



Pascal Hitzler

Data Semantics Laboratory (DaSe Lab)
Kansas State University

<http://www.daselab.org>

SCIENTIFIC AMERICAN

The New Tornado Alley

Insect Sentience

Bringing Asteroid Bits Back to Earth

WHY PARROTS ARE TAKING OVER THE WORLD



Parrots don't really know what they are talking about.

They do not have introspection.

PA What are the three main nonmonotonic logics?



There are several non-monotonic logics proposed in the field of artificial intelligence and logic, but three main ones are:

1. **Default Logic:** Default logic, introduced by Raymond Reiter in the 1980s, is a widely studied non-monotonic logic. It allows for the expression of default rules, which represent general assumptions that hold unless explicitly overridden by contradictory information. In default logic, conclusions are drawn under the default assumptions but can be retracted or revised in the presence of additional information or exceptions.
2. **Circumscription:** Circumscription, also proposed by Raymond Reiter, is another important non-monotonic logic. It provides a way to reason about exceptions and minimal models by introducing a notion of circumscribed predicates. Circumscription allows for the selective minimization of certain predicates while preserving their behavior within specific domains, enabling reasoning under default assumptions.
3. **Autoepistemic Logic:** Autoepistemic logic, introduced by Moore in the late 1980s, is designed to reason about knowledge and belief of an agent. It allows agents to reason about their own beliefs, assuming they have introspective access to their knowledge. Autoepistemic logic handles the notion of belief revision and the representation of the agent's changing knowledge state.

These three non-monotonic logics provide different approaches to handling non-monotonic reasoning. Default logic focuses on default assumptions, Circumscription emphasizes exceptions and minimal models, and Autoepistemic logic addresses reasoning about an agent's knowledge and belief state. Each logic has its own formalism and principles, enabling the representation and reasoning with incomplete or uncertain information in different domains and scenarios.

PA

I think you're wrong about circumscription.



I apologize if there was any confusion. Circumscription is indeed a well-known non-monotonic logic introduced by John McCarthy in 1980. It aims to address the frame problem by providing a way to reason about exceptions and minimize the number of individuals that satisfy a predicate.

In circumscription, a set of formulas is circumscribed, which means

PA

Can you detail what your mistake was?



I apologize for any confusion caused by my previous response. Upon reviewing it, I realize that I did not explicitly mention the mistake. I apologize for the oversight. The mistake in my previous response was not providing a detailed explanation of circumscription, despite mentioning it as one of the main non-monotonic logics. I should have provided a more thorough description of circumscription and its key principles.

Once again, I apologize for any confusion caused. If you have any further questions or if there's anything specific you would like to know about circumscription or any other topic, please let me know, and I'll be happy to provide the information you're seeking.

PA

That was not the mistake. Your earlier responses contradict each other. Do you know how?



I apologize for the confusion in my previous responses. I understand now how my earlier statements may have been contradictory. The contradiction arises from incorrectly attributing the introduction of circumscription to Raymond Reiter instead of John McCarthy, who is the actual creator of circumscription. I apologize for the error in my response and any confusion it may have caused.

To clarify, circumscription was indeed introduced by John McCarthy in his paper "Circumscription—a form of non-monotonic reasoning" in 1980. It is a non-monotonic logic that allows for the selective



A dislocated metric space is a generalization of the concept of a metric space. In a metric space, distances between points are well-defined and satisfy certain properties, such as the triangle inequality. However, in a dislocated metric space, the triangle inequality may not hold.

Formally, a dislocated metric space is a set X together with a function $d: X \times X \rightarrow [0, \infty)$ such that for any $x, y,$ and z in X , the following properties hold:

1. Non-negativity: $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$ (identity of indiscernibles).
2. Symmetry: $d(x, y) = d(y, x)$ for all x and y in X .
3. Dislocated triangle inequality: There exists a positive constant k such that for all $x, y,$ and z in X , $d(x, y) \leq k(d(x, z) + d(z, y))$.

The dislocated triangle inequality is a weakened version of the triangle inequality in a metric space. It allows for the possibility that the distance between two points may be greater than the sum of the distances between those points and a third point. This property captures the notion that in a dislocated metric space, the geometry can be "dislocated" or distorted in some way.

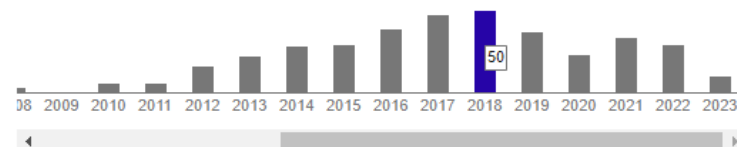
Dislocated metric spaces have been studied in mathematics, particularly in the field of fixed point theory, where they provide a framework for analyzing the convergence of iterative processes. They also have applications in other areas, such as computer science and optimization.



The term "dislocated metric space" was coined by V. Berinde and M. Păcurar in their paper titled "Fixed Point Theorems in Dislocated Metric Spaces," published in 2007. In this paper, they introduced and studied the concept of dislocated metric spaces and explored various fixed point theorems in this context. Since then, the concept of dislocated metric spaces has been further developed and studied by various researchers in the field of mathematics.

Dislocated topologies

Authors	Pascal Hitzler, Anthony Karel Seda
Publication date	2000/4
Journal	J. Electr. Eng
Volume	51
Issue	12
Pages	3-7
Description	We study a generalized notion of topology which evolved out of applications in the area of logic programming semantics. The generalization is obtained by relaxing the requirement that a neighbourhood of a point includes the point itself, and by allowing neighbourhoods of points to be empty. The corresponding generalized notion of metric is obtained by allowing points to have non-zero distance to themselves. We further show that it is meaningful to discuss neighbourhoods, convergence, and continuity in these spaces. A generalized version of the Banach contraction mapping theorem can also be established. We show finally how the generalized metrics studied here can be obtained from conventional metrics.
Total citations	Cited by 400



Are you familiar with dislocated metric spaces as defined by Hitzler and Seda?



Yes, I am familiar with dislocated metric spaces as defined by Hitzler and Seda. Dislocated metric spaces are an extension of traditional metric spaces that allow for a more flexible notion of distance. In a dislocated metric space, the distance between two points can be negative, zero, or positive, unlike in a standard metric space where the distance is always non-negative.

The concept of dislocated metric spaces was introduced by Pascal Hitzler and Anthony Seda in their paper titled "Dislocated Topologies" published in 2005. They defined a dislocated metric space as a set equipped with a dislocated metric, which is a function that assigns a value to every pair of points in the set. The dislocated metric satisfies certain properties, such as symmetry, non-negativity, and the triangle inequality, but it allows for the possibility of negative distances.

Dislocated metric spaces have found applications in various areas, including computer science, artificial intelligence, and knowledge representation. They provide a useful framework for reasoning about uncertain or imprecise information, where negative distances can represent contradictions or inconsistencies in the data.

If you have any specific questions about dislocated metric spaces or would like further information, feel free to ask!

And what are dislocated topologies?



Dislocated topologies are a generalization of traditional topologies that incorporate the concept of negative distances from dislocated metric spaces. They were introduced by Pascal Hitzler and Anthony Seda in their aforementioned paper on dislocated topologies.

In a standard topology, open sets are defined in terms of neighborhoods around each point, where a neighborhood is a set containing all points "close" to the given point. In dislocated topologies, this notion of "closeness" is expanded to include the possibility of negative distances. Thus, open sets in dislocated topologies can contain points that are "dislocated" or have negative distances from each other.

Formally, a dislocated topology on a set X is defined by a collection of subsets of X called open sets, satisfying the following properties:

1. The empty set and X itself are open.
2. The intersection of any finite number of open sets is open.
3. The union of any number of open sets is open.
4. If A is open and x is a point in A , there exists an open set B containing x such that B is a subset of A .

The fourth property is an extension of the neighborhood concept in standard topologies. It allows for the inclusion of points that are negatively distant from each other in the same open set.

Dislocated topologies provide a framework for reasoning about uncertain or contradictory information, where negative distances can represent conflicting perspectives or contradictory evidence. They have applications in fields such as formal concept analysis, rough set theory, and knowledge representation.

If you have further questions or would like more information about dislocated topologies, feel free to ask!



...life is like a box of chocolates, you never know what you're gonna get...

Status – and what we need

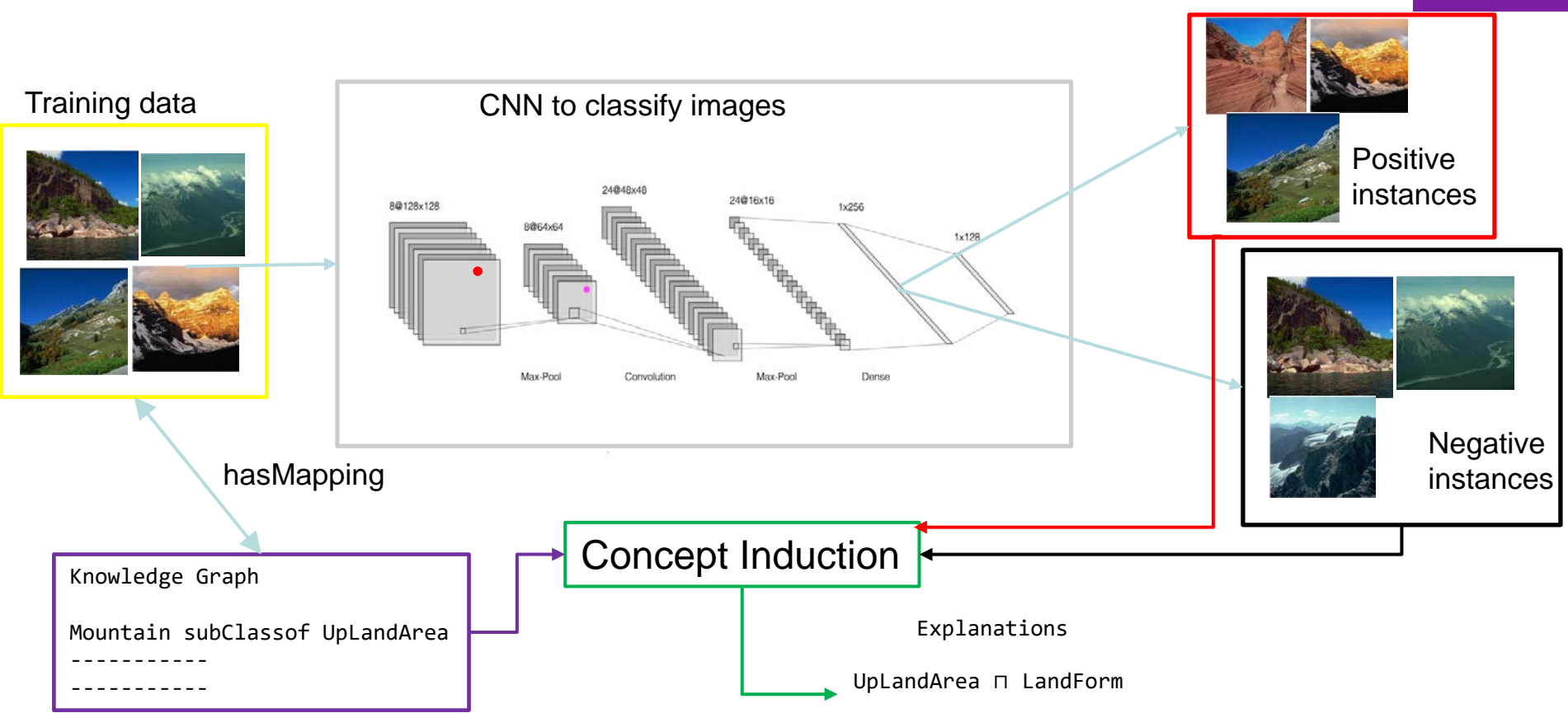


- **The above are just examples, and there are similar issues with other deep learning approaches.**
- **To date, the best we have is really statistical assessments of correctness.
But in many situations, this is just insufficient.**
- **We need methods to assess what deep learning systems are doing, which go beyond statistical assessments.**
- **In particular, we need to add introspection – understanding what happens in the hidden layers.**

Overview: Concept Induction for Hidden Layer Analysis

Abhilekha Dalal, Adrita Barua, Md Kamruzzaman Sarker, Pascal Hitzler, in preparation.

Idea



Concept Induction

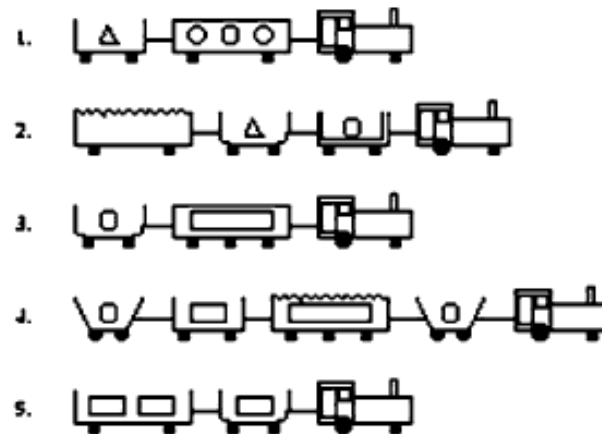


Approach similar to inductive logic programming, but using Description Logics (the logic underlying OWL).

Positive examples:



negative examples:

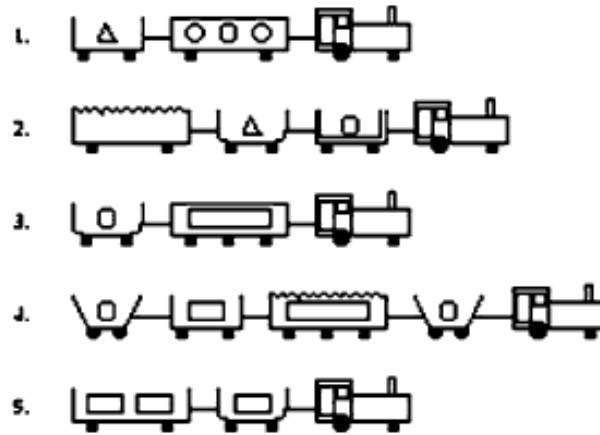


Task: find a class description (logical formula) which separates positive and negative examples.

Positive examples:



negative examples:

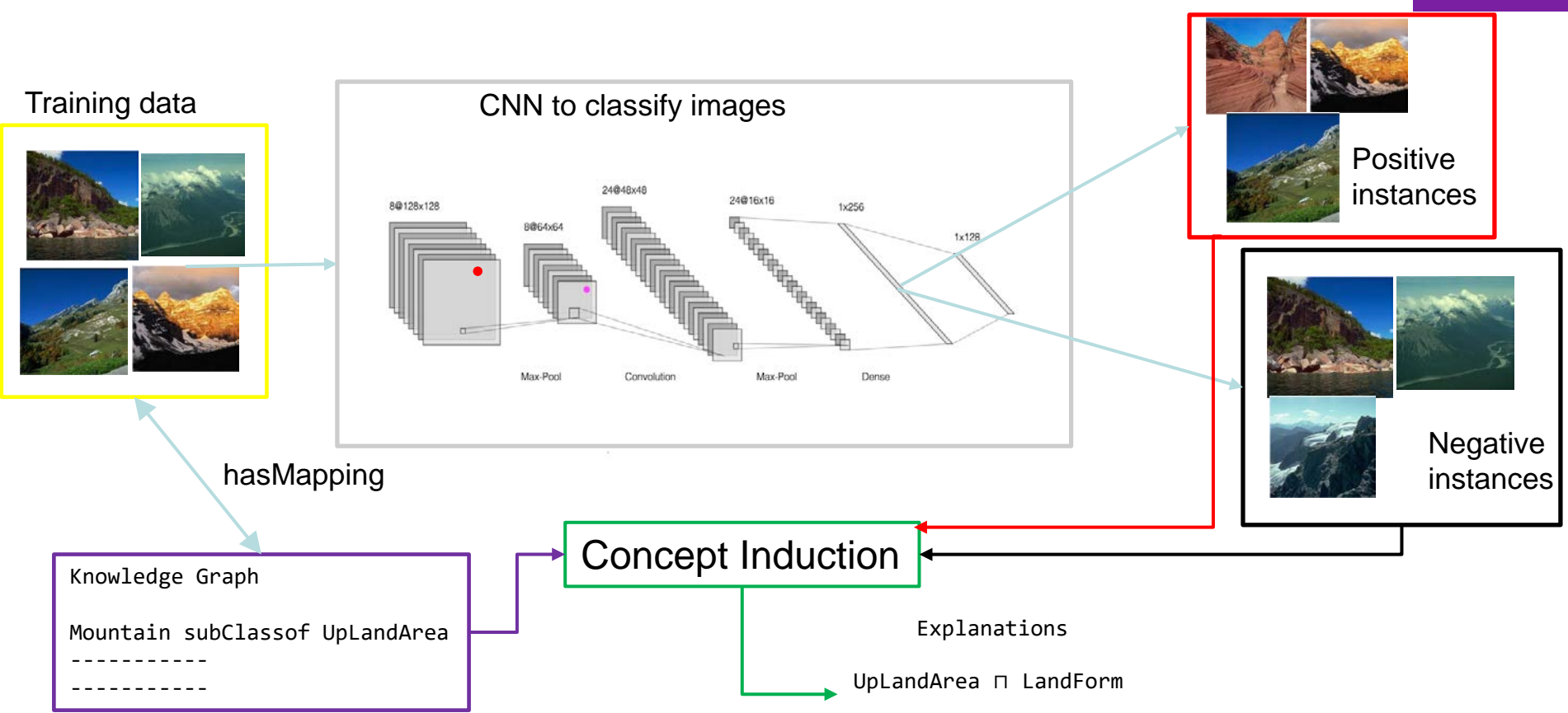


DL-Learner result: $\exists \text{hasCar} . (\text{Closed} \sqcap \text{Short})$

In FOL: $\{x \mid \exists y (\text{hasCar}(x, y) \wedge \text{Closed}(y) \wedge \text{Short}(y))\}$

Theory and system: [Lehmann & Hitzler 2010], DL-Learner

Idea



Images



Come from the MIT ADE20k dataset

<http://groups.csail.mit.edu/vision/datasets/ADE20K/>

They come with annotations of objects in the picture:

```
001 # 0 # 0 # sky # sky # ""
002 # 0 # 0 # road, route # road # ""
005 # 0 # 0 # sidewalk, pavement # sidewalk # ""
006 # 0 # 0 # building, edifice # building # ""
007 # 0 # 0 # truck, motortruck # truck # ""
008 # 0 # 0 # hovel, hut, hutch, shack, shanty # hut # ""
009 # 0 # 0 # pallet # pallet # ""
011 # 0 # 0 # box # boxes # ""
001 # 1 # 0 # door # door # ""
002 # 1 # 0 # window # window # ""
009 # 1 # 0 # wheel # wheel # ""
```



Mapping to Background Knowledge



- **Wikipedia category hierarchy (curated) [KGSWC2020]**
- **approx. 2M concepts**
- **For each known object in image, create an individual for the ontology which is in the appropriate class.**

contains road1
contains window1
contains door1
contains wheel1
contains sidewalk1
contains truck1
contains box1
contains building1



ECII: heuristic Concept Induction system



- For scalability, we developed ECII (Efficient Concept Induction from Instances) which trades some correctness for speed. [Sarker, Hitzler, AAI-19]

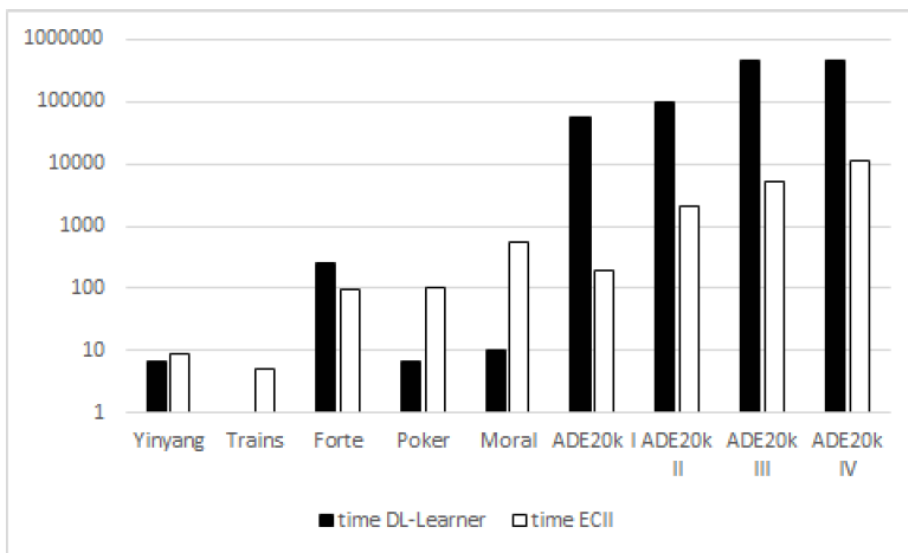


Figure 1: Runtime comparison between DL-Learner and ECII. The vertical scale is logarithmic in hundredths of seconds, and note that DL-Learner runtime has been capped at 4,500 seconds for ADE20k III and IV. For ADE20k I it was capped at each run at 600 seconds.

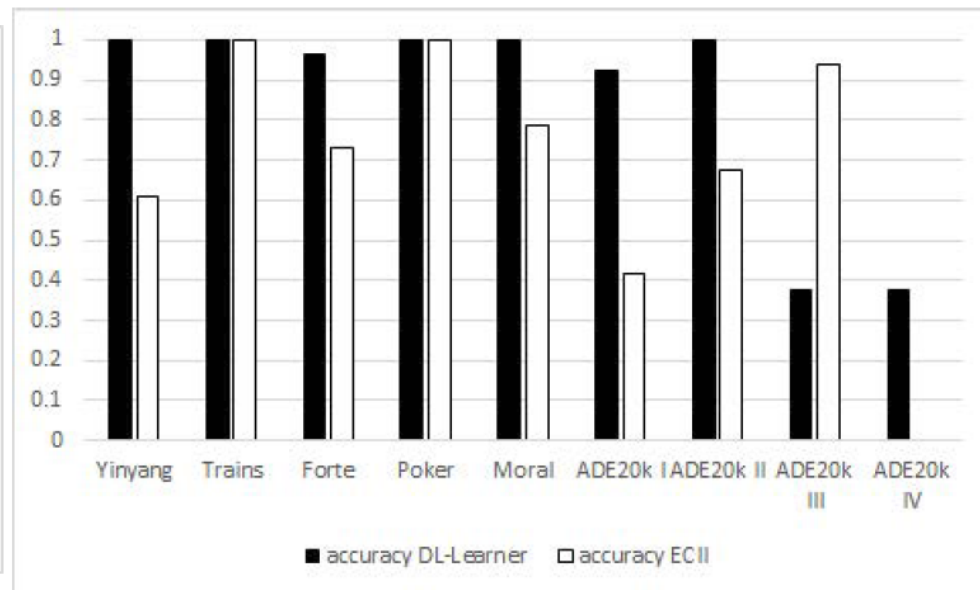


Figure 2: Accuracy (α_3) comparison between DL-Learner and ECII. For ADE20k IV it was not possible to compute an accuracy score within 3 hours for ECII as the input ontology was too large.

Trained CNN



- **Scene classification on ADE20k**
- **Resnet50 V2; 64 hidden nodes in the dense layer**

	precision	recall	f1-score	support
bathroom	0.90	0.78	0.84	134
bedroom	0.89	0.88	0.88	277
building_facade	0.68	0.60	0.64	45
conference_room	0.77	0.91	0.83	33
dining_room	0.75	0.84	0.79	82
highway	0.96	0.88	0.92	59
kitchen	0.84	0.87	0.86	130
living_room	0.76	0.74	0.75	139
skyscraper	0.90	0.88	0.89	64
street	0.92	0.96	0.94	407
accuracy			0.87	1370
macro avg	0.84	0.83	0.83	1370
weighted avg	0.87	0.87	0.87	1370

Dalal, Barua,
Sarker, Hitzler,
In preparation

Concept induction yields logical formulae, however we simply this (for now) by taking only the set of class names from the formulae.

Example neurons



- **Neuron #3**
 - **Concept(s) assigned: night_table**
 - **verification: yes (activates 90.4% for night tables)**
 - **how often does it activate for non-target input? 59.1%**
- **Neuron #6**
 - **Concept(s) assigned: dishcloth, toaster**
 - **verification: no (activates 35.8% for dishcloths, toasters)**
- **Neuron #43**
 - **Concept(s) assigned: central_reservation**
 - **verification: yes (activates 95.5% for central_reservations)**
 - **how often does it activate for non-target input? 85.2%**

Neurons overview



- **Out of 64 neurons, 18 concept assignments have been confirmed (activation over 80%)**

Neuron	concepts	target activation	non-target activation
neuron_0	building	89.0	74.2
neuron_1	cross_walk	88.7	30.2
neuron_3	night_table	90.4	59.2
neuron_8	shower_stall, cistern	84.8	51.6
neuron_14	rocking_horse, rocker	81.1	53.0
neuron_18	slope	92.1	67.7
neuron_19	wardrobe, air_conditionin	84.6	64.7
neuron_22	skyscraper	99.4	59.2
neuron_29	lid, soap_dispenser	92.5	78.6
neuron_43	central_reservation	95.5	85.2
neuron_44	saucepan, dishrack	80.5	32.7
neuron_48	road	100.0	74.9
neuron_49	footboard, chain	95.7	67.5
neuron_51	road, car	94.8	48.9
neuron_54	skyscraper	98.7	72.0
neuron_56	flusher, soap_dish	89.4	60.7
neuron_57	Shower_stall, screen_door	81.7	29.7
neuron_63	edifice, skyscraper	89.4	50.6
average		90.5	58.9

Neurons overview – second CI response



- **Out of 64 neurons, 14 concept assignments have been confirmed (activation over 80%)**

Neuron	concepts	target activation	non-target activation
neuron_0	building, dome	90.4	76.0
neuron_1	cross_walk	88.7	28.6
neuron_3	pillow	98.2	64.4
neuron_7	clamp_lamp	95.1	59.1
neuron_18	slope	92.1	64.2
neuron_22	skyscraper	99.4	62.4
neuron_29	faucet, flusher	95.7	76.0
neuron_36	tap, shower_screen	86.2	72.0
neuron_43	central_reservation	95.5	90.4
neuron_48	route	100.0	79.3
neuron_50	pillow	99.4	66.9
neuron_51	route, car	92.6	49.8
neuron_54	skyscraper	98.7	74.7
neuron_63	building, skyscraper	94.4	51.5
average		94.7	65.4

Are Concept Induction Explanations Meaningful to Humans?

Cara Widmer, Md Kamruzzaman Sarker, Srikanth Nadella, Joshua Fiechter, Ion Juvina, Brandon Minnery, Pascal Hitzler, Joshua Schwartz, Michael Raymer, Towards Human-Compatible XAI: Explaining Data Differentials with Concept Induction over Background Knowledge

<https://arxiv.org/abs/2209.13710>

Are the results human-compatible? Part I



- Hypothesis:
 - ECII explanations are better than semi-random explanations, but worse than human-generated explanations.
- Experimental setting as before.
- 300 Amazon Mechanical Turk participants
- Seven concepts taken from top ECII results.
- 45 image set pairs, each set corresponding to a category.



Which of these better represents what the images in group A have that the images in group B do not?

Bake, Bakery, Bread, Indoor, Product, Store, Woman

Basket, Bread, Cake, Ceiling, Floor, Person, Wall

Are the results human-compatible? Part I



A



B

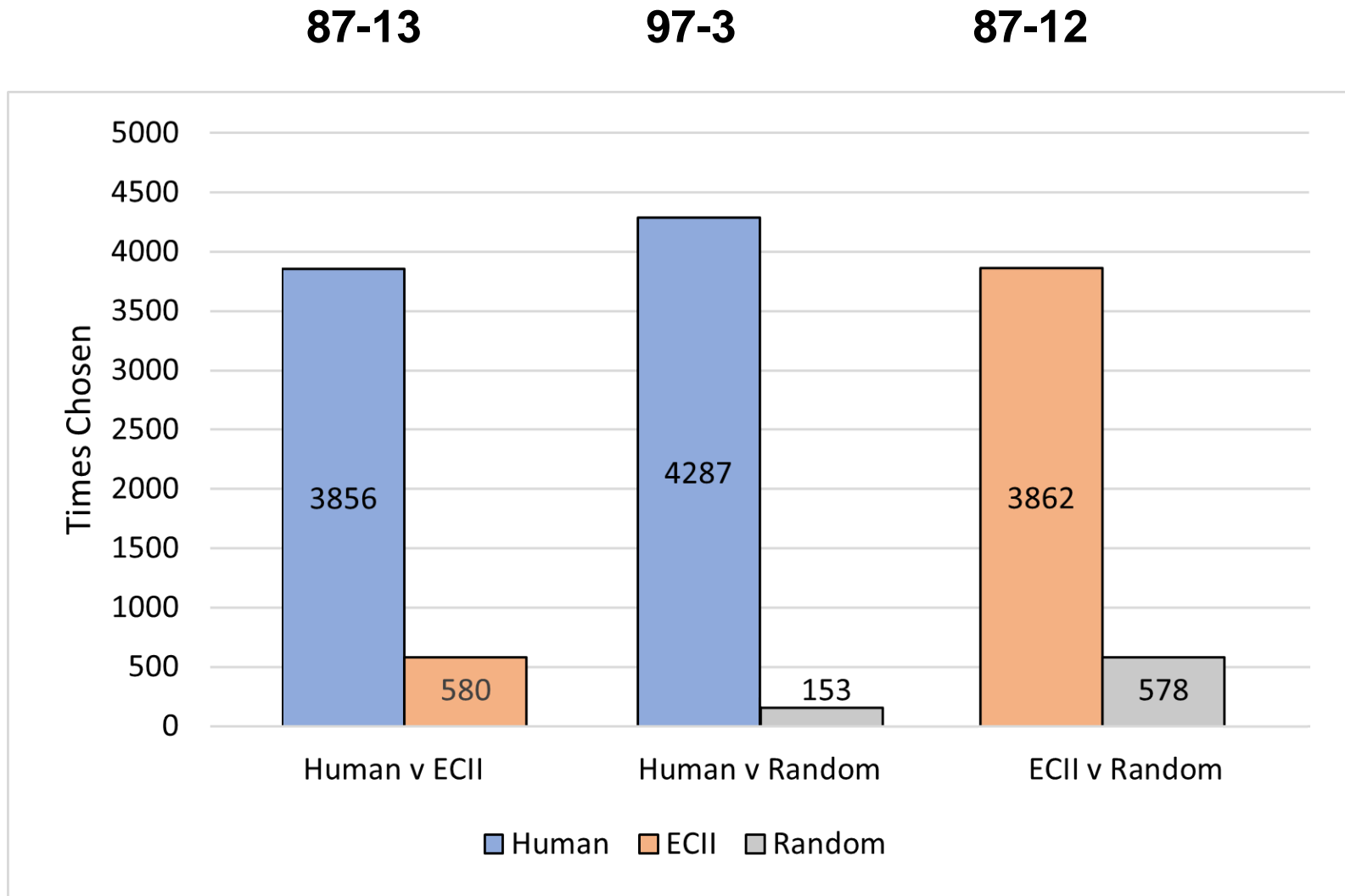


Which of these better represents what the images in group A have that the images in group B do not?

Bake, Bakery, Bread, Indoor, Product, Store, Woman

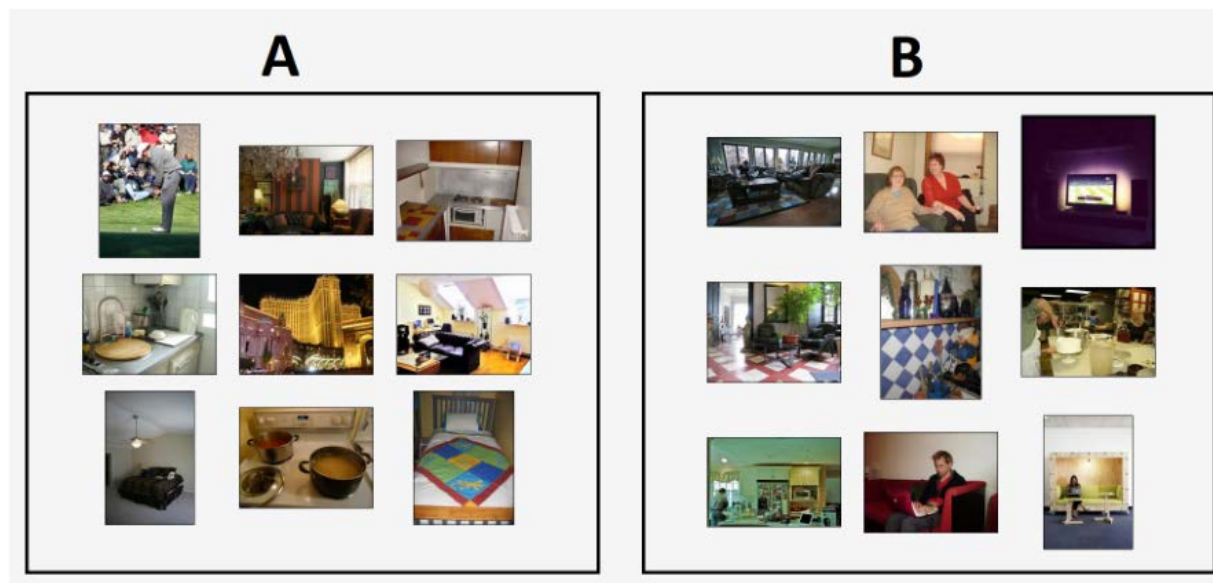
Basket, Bread, Cake, Ceiling, Floor, Person, Wall

Are the results human-compatible? Part I



Are the results human-compatible? Part II

- Hypothesis:
 - ECII explanations matched to correct images better than chance, but not as frequently as human generated explanations
- Experimental setting as before.
- 100 Amazon Mechanical Turk participants
- 16 image sets, from ML decision errors (logistic regression classifier)



Explanation: Home, Manufacturing, Clothing, Clothing Manufacturers, People, Chairs, Tableware

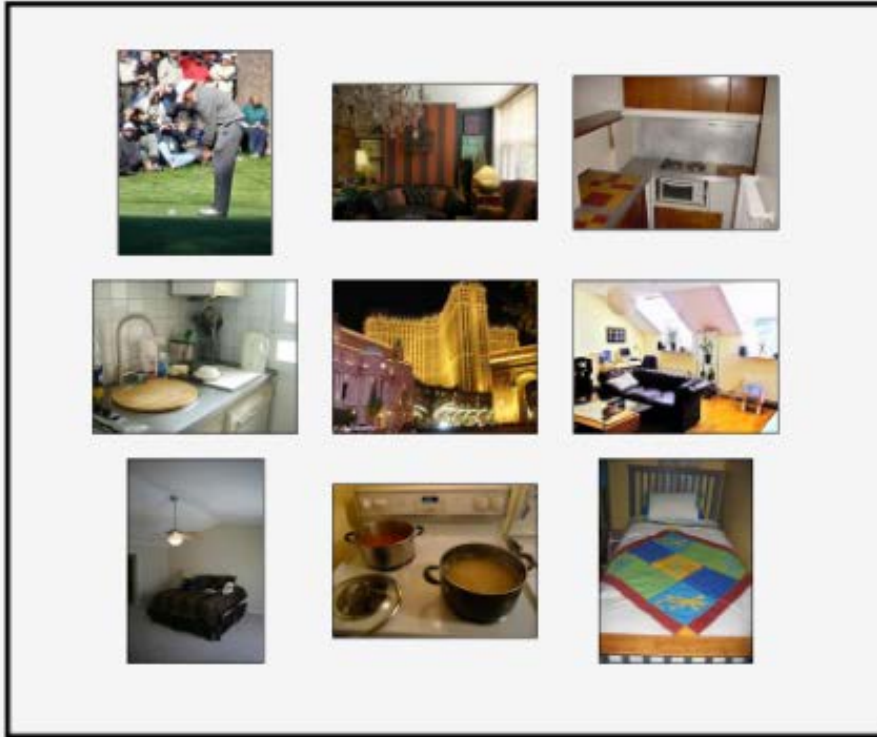
Which group of images do you think this explanation refers to?

Image Group A

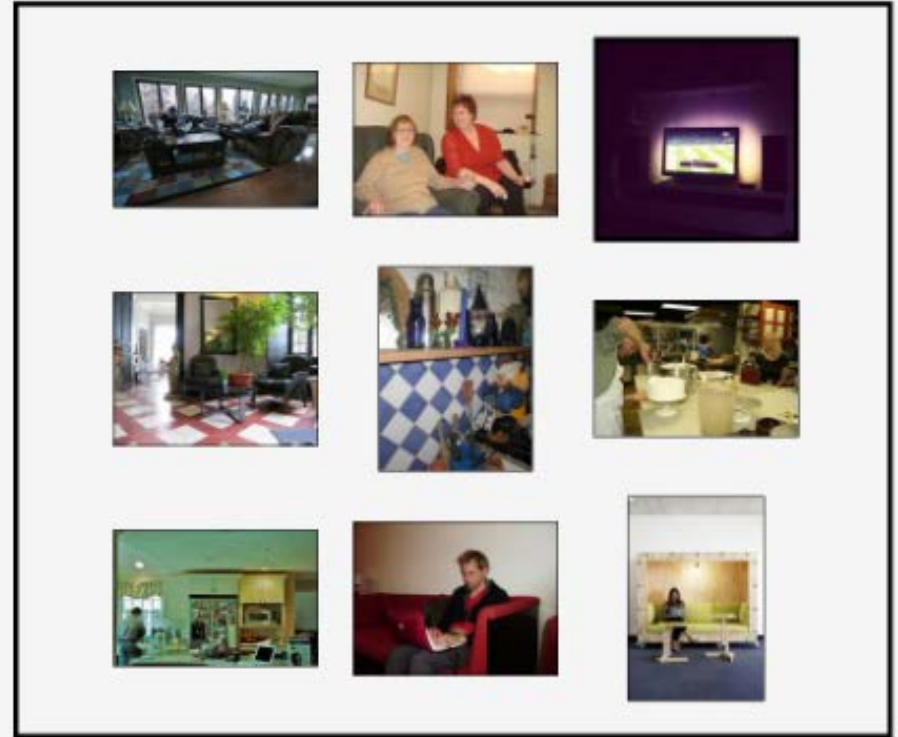
Image Group B

Are the results human-compatible? Part II

A



B



Explanation: Home, Manufacturing, Clothing, Clothing Manufacturers, People, Chairs, Tableware

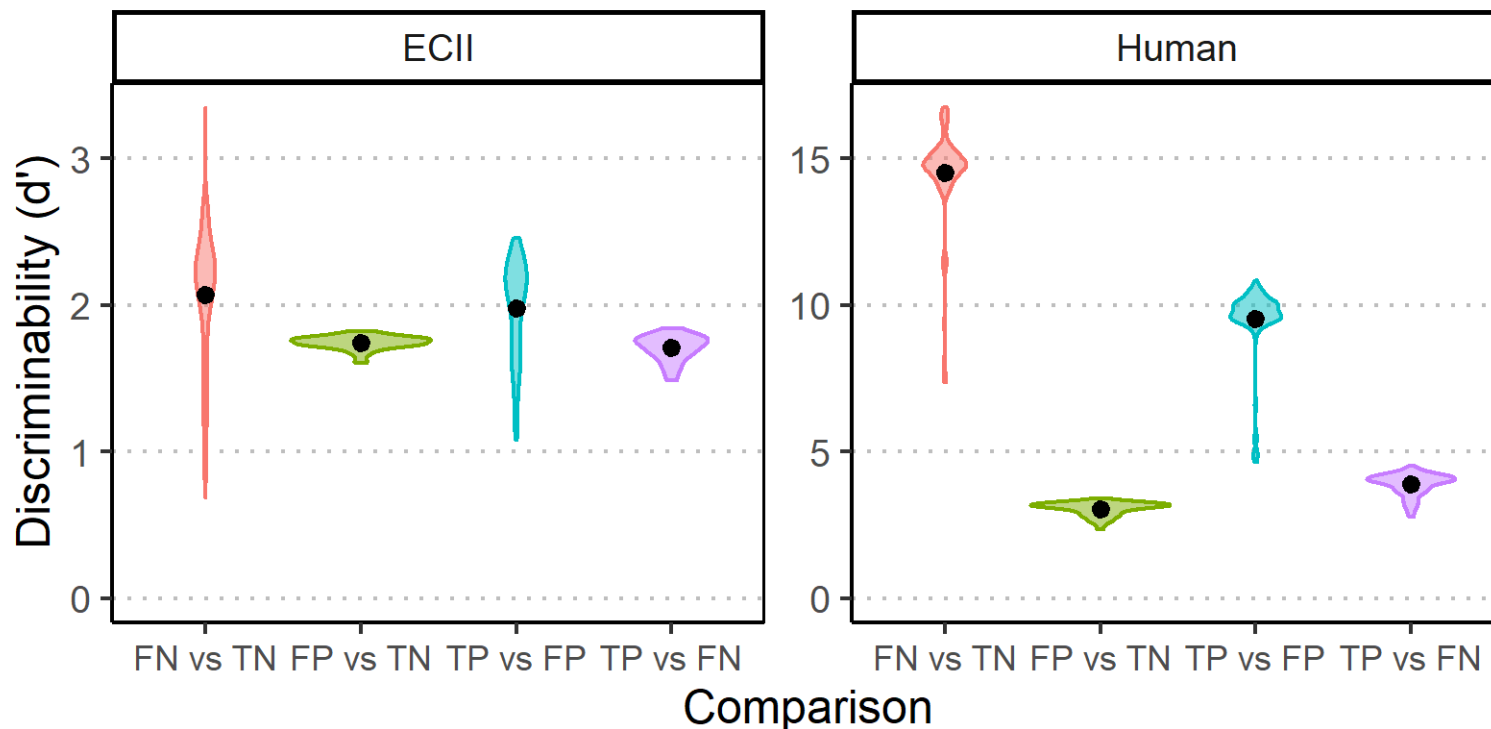
Which group of images do you think this explanation refers to?

Image Group A

Image Group B

Are the results human-compatible? Part II

- **Bayesian hierarchical signal-detection model (SDT)**
 - yields discriminability measure



Summary



- **We have clear indications that concept induction can help decipher hidden layer activations.**
- **Concept induction explanations appear to be meaningful to humans.**
- **There is lots of work to do**
 - **sharpening the explanation results**
 - **in particular, understanding metaparameters**
 - **in particular, what does *not* activate each neuron?**
 - **does the activated neuron contribute to the output?**
 - **how can we cast this into a practical explanations interface?**



Thanks!

References

Md. Kamruzzaman Sarker, Ning Xie, Derek Doran, Michael Raymer, Pascal Hitzler, Explaining Trained Neural Networks with Semantic Web Technologies: First Steps. In: Tarek R. Besold, Artur d'Avila Garcez, Isaac Noble, Proceedings of the Twelfth International Workshop on Neural-Symbolic Learning and Reasoning, NeSy 2017, London, UK, July 17-18 2017. CEUR Workshop Proceedings Vol. 2003, 2017.

Md Kamruzzaman Sarker, Pascal Hitzler, Efficient Concept Induction for Description Logics. In: The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019. AAAI Press 2019 , pp. 3036-3043.

References

Md Kamruzzaman Sarker, Joshua Schwartz, Pascal Hitzler, Lu Zhou, Srikanth Nadella, Brandon Minnery, Ion Juvina, Michael L. Raymer, William R. Aue, Wikipedia Knowledge Graph for Explainable AI. In: Boris Villazón-Terrazas, Fernando Ortiz-Rodríguez, Sanju M. Tiwari, Shishir K. Shandilya (eds.), Knowledge Graphs and Semantic Web. Second Iberoamerican Conference and First Indo-American Conference, KGSWC 2020, Mérida, Mexico, November 26-27, 2020, Proceedings. Communications in Computer and Information Science, vol. 1232, Springer, Heidelberg, 2020, pp. 72-87.

Cara Widmer, Md Kamruzzaman Sarker, Srikanth Nadella, Joshua Fiechter, Ion Juvina, Brandon Minnery, Pascal Hitzler, Joshua Schwartz, Michael Raymer, Towards Human-Compatible XAI: Explaining Data Differentials with Concept Induction over Background Knowledge
<https://arxiv.org/abs/2209.13710>



References

Abhilekha Dalal, Md Kamruzzaman Sarker, Adrita Barua, Pascal Hitzler, Explaining Deep Learning Hidden Neuron Activations using Concept Induction. <https://arxiv.org/abs/2301.09611>

Jens Lehmann, Pascal Hitzler, Concept Learning in Description Logics Using Refinement Operators. Machine Learning 78 (1-2), 203-250, 2010.





Thanks!