LOD 2020, Siena, Italy
23 July 2020

Neural-Symbolic Computing for Trusted AI

Professor Artur d’Avila Garcez, FBCS
City, University of London
a.garcez@city.ac.uk
@AvilaGarcez
http://www.staff.city.ac.uk/~aag/

Outline

- Reconciling Symbolic AI and Deep Learning: Logic and Learning in Neural Nets (ML + KR)
- Neural-Symbolic Methodology
- Explainable AI
- Impact case in industry
- Challenges and future directions

Neural-Symbolic AI in the 2020s

Debate between Yoshua Bengio and Gary Marcus, Dec 2019: https://www.youtube.com/watch?v=EeqwFjqFvJA

AAAI 2020 – A turning point?

- Debate between 2019 Turing award winners Hinton, Bengio and LeCun and Nobel laureate Dan Kahneman (Thinking Fast and Slow): “System 1 certainly knows language… System 2 does involve certain manipulation of symbols”
- Robert S. Engelmore Memorial Award Lecture by Henry Kautz: Introduced a taxonomy for Neural-Symbolic AI
- Director of MIT IBM Watson AI lab David Cox's keynote: AutoAI: infusing knowledge into NLP networks, Argumentation, XAI, Causality, Planning.
- At recent top AI conference: % of papers with “learning” (30%), “reasoning” (25%), “learning” AND “reasoning” (2%), but increasing quickly c.f. AAAI 2020, KR+ML at NeurIPS and KR, IJCLR 2020/21 + ML journal track on Learning and Reasoning

Latest AI revolution mainly due to Deep Learning

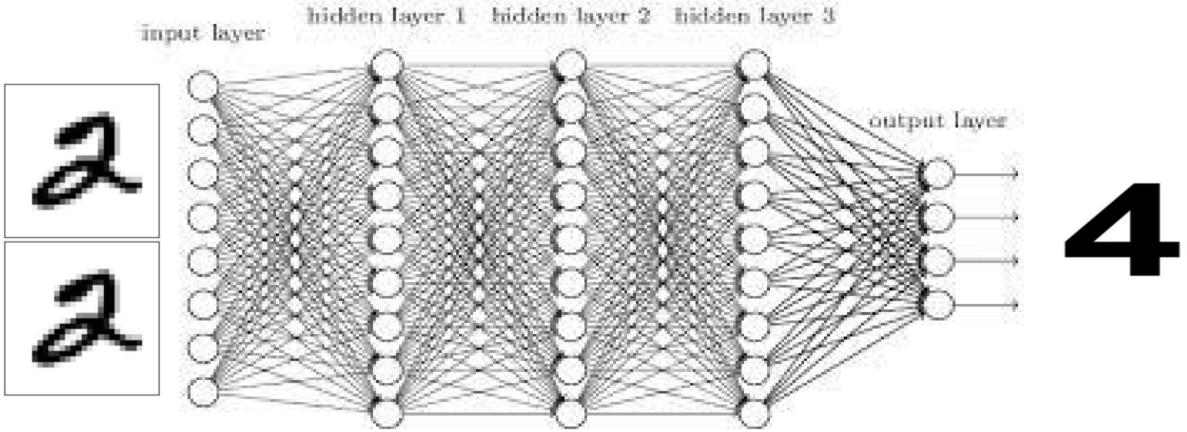
A very nice original idea (deep belief nets; semi-supervised learning) later engineered into large neural systems that work well in practice using *backprop.*

State-of-the-art at computer vision tasks, speech, audio and games, language translation

With good progress at Question Answering and multi-modal learning

Problems with Deep Learning

- Limited at reasoning (propositional modal logic)
- Lacks relational learning (FOL, StaR-AI)
- Fails at Extrapolation (adversarial attacks)




```
add(X,Y,Z):-
  Z is X+Y.
```

345+223 = ?

Neural-Symbolic Computing (1)

- Neural networks provide the machinery for effective learning and computation
- But perception alone is insufficient: AI requires reasoning, explanation, planning, transfer...
- Need for rich knowledge representation: non-monotonic, relational (with variables), recursion, time and uncertainty

Neural-symbolic system = (modular) neural networks with logical (compositional) interpretation

Neural-Symbolic Computing (2)

Symbolic

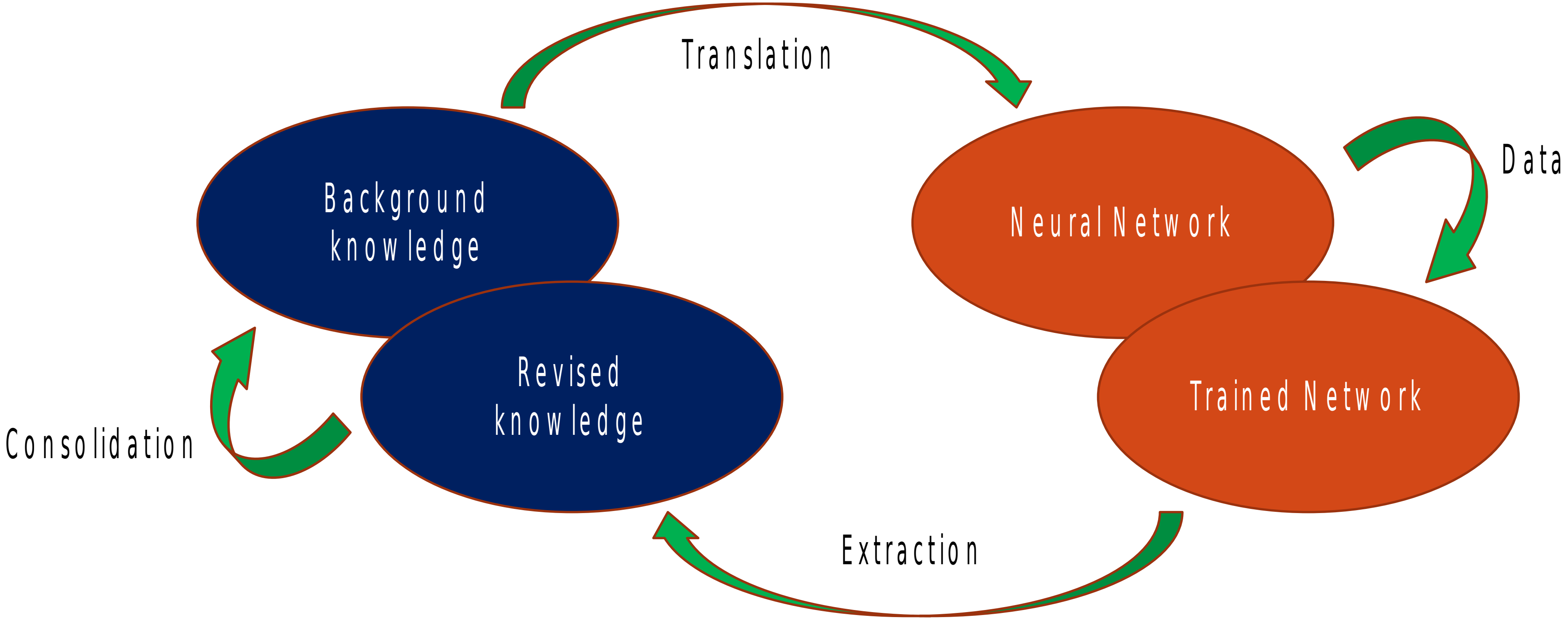
propositional logic
temporal logic
logic programming
nonmonotonic logic
propositional modal logic
first-order logic
higher-order logic

Neural

CNN, RNN, GNN,
LSTM, GANs,
Transformer...

A well-established hierarchy by contrast with scarce results on expressiveness; neural-symbolic computing can establish correspondence proofs and help identify the fundamental ingredients of neural computation

Neural-Symbolic Learning Cycle



Neural-Symbolic Methodology

high-level symbolic representations
(abstraction, recursion, relations, modalities)

translations

low level, efficient neural structures
(with the same, simple architecture throughout)

Analogy with programming languages: AI compiler

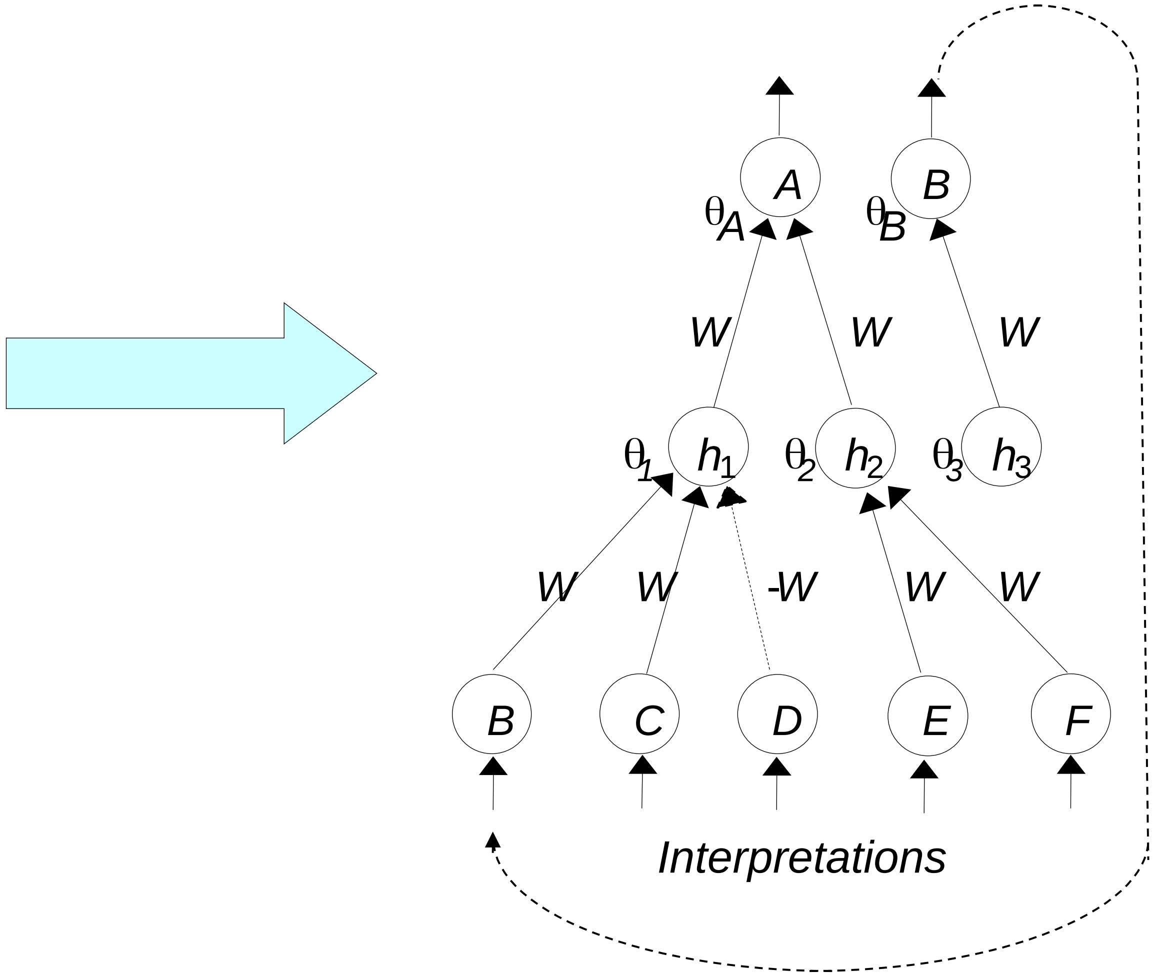
(dealing with multiple levels of abstraction)

Connectionist Inductive Logic Programming (CILP)

r_1: $A \leftarrow B, C, \sim D$;

r_2 : $A \leftarrow E, F$;

r_3 : $B \leftarrow$



THEOREM: For any logic program P there exists a neural network N such that N computes P

Applications

Training and assessment in simulators,
Robocup,
Evolution of software models,
Bioinformatics,
Power plant fault diagnosis,
Semantic web, Ontology learning,
General game playing,
Visual intelligence,
Consumer protection in gambling,
Fraud prevention,
Investment decisions, etc.

Power Plant Fault Diagnosis

First real-world application of CILP

Rules and Examples

Background Knowledge (35 rules with errors)

278 examples of single and multiple faults

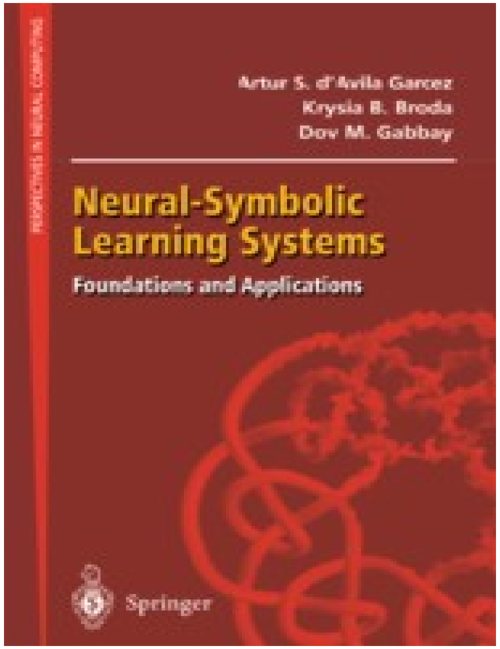
Fault(ground,close-up,line01,no-bypass) IF
Alarm(instantaneous,line01) AND
Alarm(ground,line01)

There is a fault at transmission line 01, close to the power plant generator, due to an over-current in the ground line of transmission line 01, which occurred when the system was not using the bypass

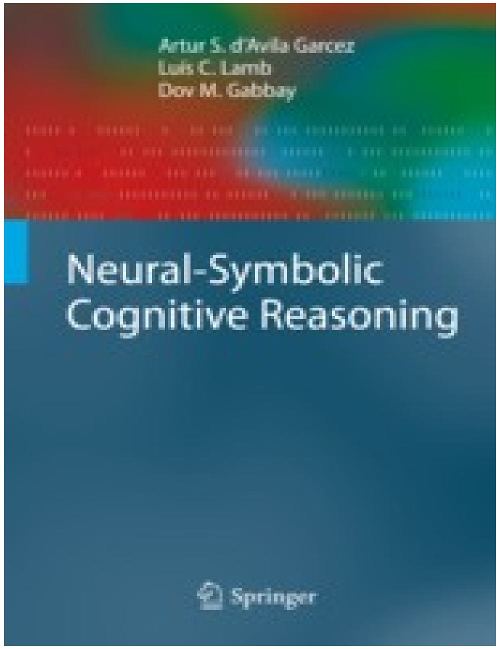
Theory and Applications

For details:

d'Avila Garcez, Broda and Gabbay, Neural-Symbolic Learning Systems, Springer, 2002.



d'Avila Garcez, Lamb and Gabbay, Neural-Symbolic Cognitive Reasoning, Springer, 2009.



Knowledge Extraction is an integral part of neural-symbolic methodology

- Soundness implies high fidelity
- Explanation via proof history (goal-directed reasoning)
- But it needs to take place at the right level of abstraction (probably not pixel level)
- Added benefits:

  Knowledge-based Transfer Learning

  System maintenance/improvement

Lots of recent work on XAI

- In a nutshell:

  Feed-forward networks => trees

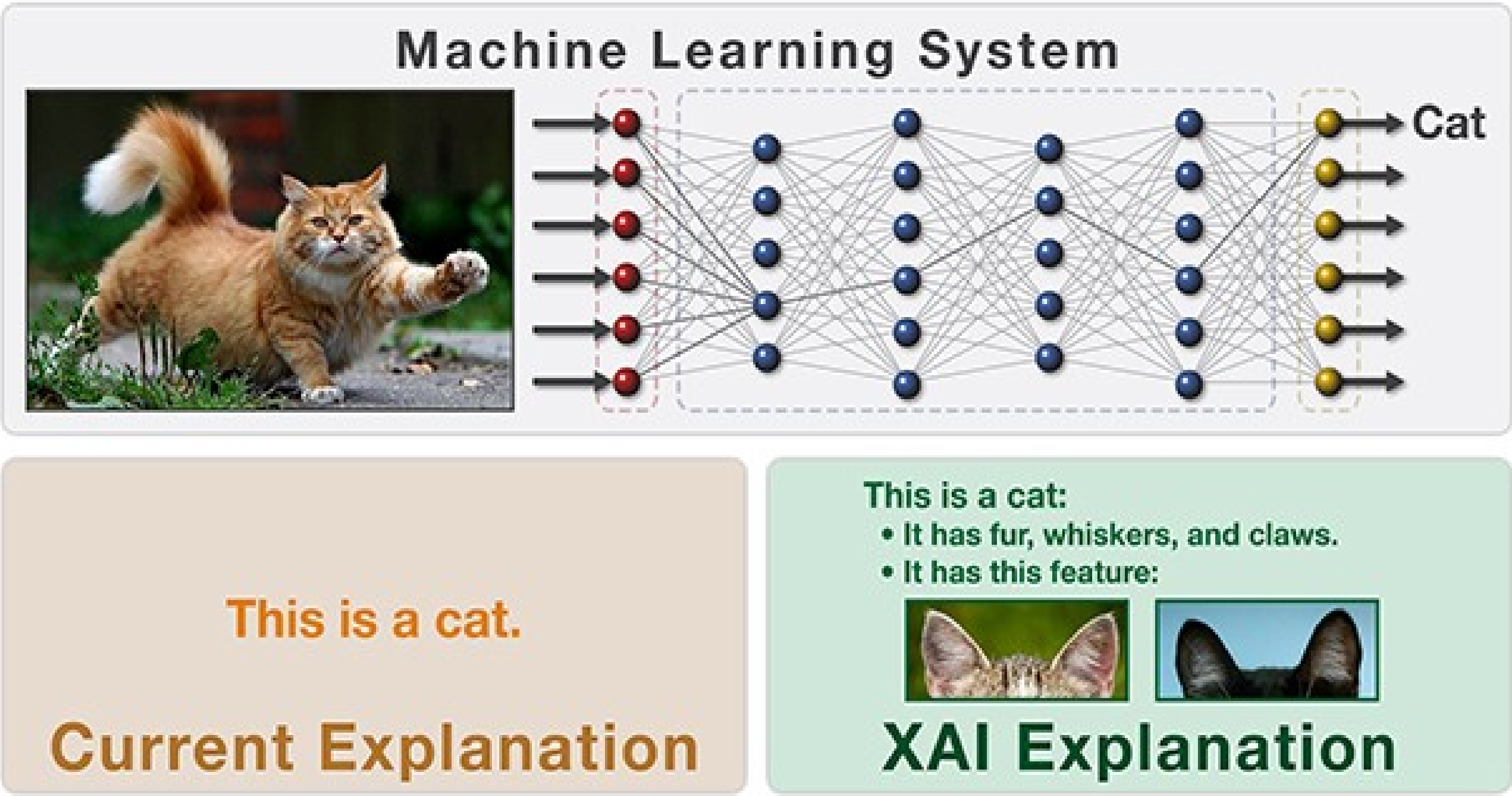
  Recurrent networks => graphs
- Knowledge Extraction from RBMs, Deep Belief Nets, Random Forests
- Layer-wise Knowledge Extraction from CNNs
- Counterfactual Local Explanations

Learning and Representing Temporal Knowledge in Recurrent Networks, Rafael V. Borges, Artur d'Avila Garcez, Luis C. Lamb, IEEE TNNLS, 2011

Extracting M of N Rules from Restricted Boltzmann Machines, Simon Odense and Artur d'Avila Garcez, ICANN 2017

Measurable Counterfactual Local Explanations for Any Classifier. Adam White, Artur d'Avila Garcez, https://arxiv.org/abs/1908.03020, To appear at ECAI 2020

DARPA Explainable AI (XAI) programme

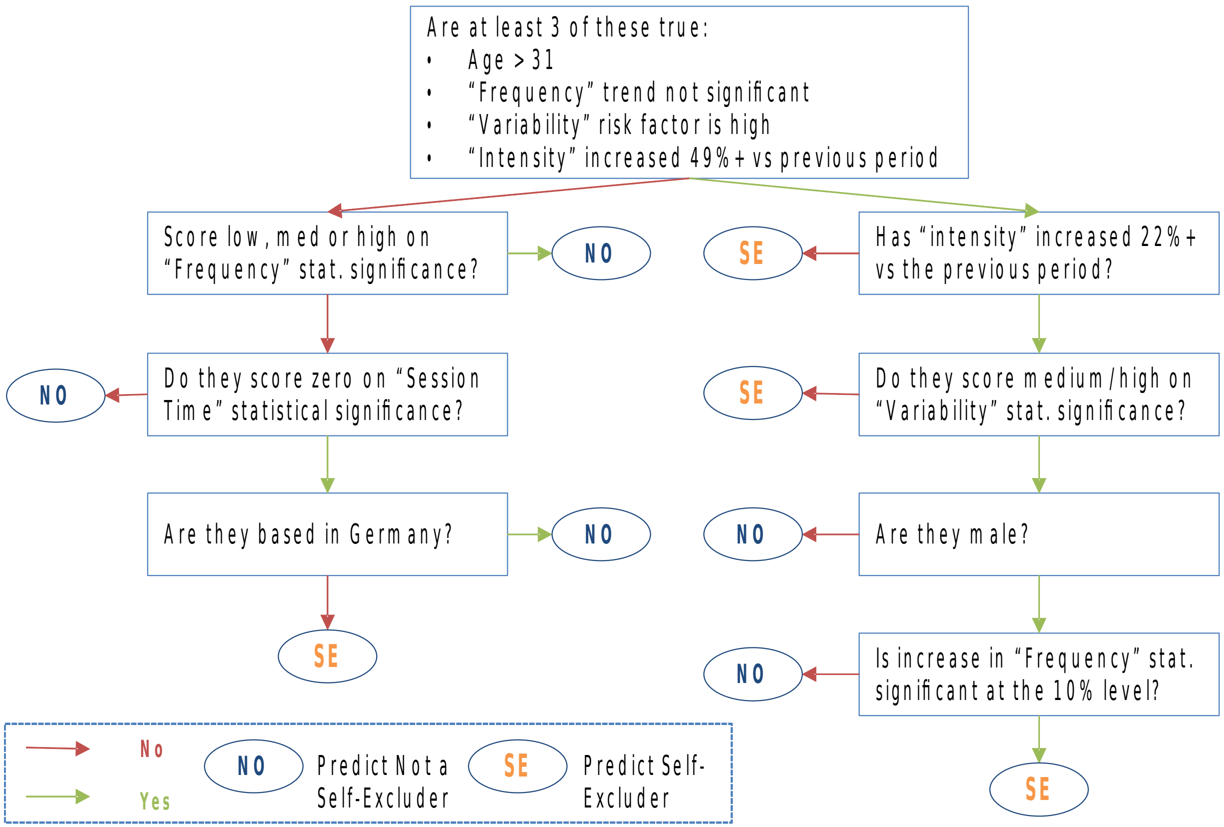


- XAI = Interpretable ML
- Explanation = knowledge extraction, not XAI

Recent industry application: Reducing harm from gambling

- EPSRC/InnovateUK project with BetBuddy Ltd, later acquired by Playtech plc
- NN and RF trained to predict harm (with self-exclusion as a proxy) given online gambling transaction data: frequency of play, betting intensity, night time play, etc.
- ML system is required to provide explanations to: the regulator, the gambling operator, Playtech’s data scientist and the customer/player

Decision Tree extracted from Neural Network:



Extracting tree-structured representations of trained networks, Mark W. Craven and Jude W. Shavlik, NIPS 1995

C. Percy, A. S. d'Avila Garcez, S. Dragicevic, M. Franca, G. Slabaugh and T. Weyde. The Need for Knowledge Extraction: Understanding Harmful Gambling Behavior with Neural Networks, In Proc. ECAI 2016, The Hague, September 2016.

Frosst and Hinton: Distilling a Neural Network Into a Soft Decision Tree, AI-IA CEX, Bari, Sep 2017.

Logic Tensor Network (LTN)

- A framework where the elements of a first-order logic signature are grounded onto a distributed representation, vectors and tensors, for learning with neural networks
- FOL statements are translated into real-valued constraints to be added to the loss function by taking a many-valued (fuzzy logic) interpretation
- The user can interact with the system by asking questions/querying the data-driven network on its understanding of high-level knowledge

Ivan Donadello, Luciano Serafini, Artur d'Avila Garcez. Logic Tensor Networks for Semantic Image Interpretation. In Proc IJCAI 2017.

Current research

- Logic Tensor Networks with Attention applied to Reasoning with Images, Videos and NLP
- Applications of neural-symbolic AI and XAI in healthcare (OCT scans and COVID19 x-ray and blood) and business/finance (investment decisions)
- Semi-supervised transfer learning: Sequence Classification Restricted Boltzmann Machines With Gated Units. S. Tran, A. Garcez, T Weyde, J Yin, Q Zhang, M Karunanithi, IEEE TNNLS, 2020
- Neuro-Symbolic Probabilistic Argumentation Machines: Regis Riveret, Son Tran and Artur Garcez, To appear KR2020
- Efficient learning from fewer data: CILP++ and Graph Nets. Graph Neural Networks Meet Neural-Symbolic Computing: A Survey and Perspective. Luis C. Lamb, Artur Garcez, Marco Gori, Marcelo Prates, Pedro Avelar, Moshe Vardi, IJCAI 2020
- Ethics of AI: Principles and audit processes for the gambling Industry (consultancy)

Research Challenges

- First-order logic knowledge extraction from very large networks (sound and efficient, explaining entire model)
- Goal-directed (commonsense or combinatorial) reasoning about what has been learned by a deep network
- Human-network communication; an agent that asks questions and checks its understanding!
- Necessary and sufficient ingredients for neural-symbolic AI
- Evaluation of representations, models and approaches

Contributions to neural-symbolic computing in each of the above areas may become seminal and help define the area

Conclusion: Why Neurons and Symbols

To study the statistical nature of learning and the logical nature of reasoning.

To provide a unifying foundation for robust learning and efficient reasoning.

To develop effective computational systems for integrated reasoning and learning.

Figure 1. Conflict between theoretical extremes.

Thank you!