Human-Allied Artificial Intelligence



Sriraam Natarajan

The University of Texas at Dallas











Not cally.



Facts not fiction $\ensuremath{\textcircled{\odot}}$



Source: National Geographic











Human-Allied Al



Can we build systems that can seamlessly **interact** with, **learn** from, and **collaborate** with humans?

Human-Allied AI: The Assistant





Human-Allied Al: The Apprentice

Human to machine: Please complete this task!

Example: "Automate physician reports!" or "Enter this data into the electronic health record!"



Human-Allied AI: The Collaborator



Al, according to the world: take your data spreadsheet...





IEEE CVT Dallas, October 15, 2019

40

Unfortunately, in reality...



Challenges to HAAI

Different types and formats of data

Different scales of data

Different **frequencies** of data streams

Noise in measurements/sensors/data collection

Changes in acquired knowledge

Uncertain side-effects of actions

Partial observability of the world

Long-term effects of decision-making





The Most Important Challenge?

Humans!!!!

Humans reason **approximately** Humans act **unpredictably**



* Thanks to Rao Khambampati

Understanding a human model is crucial

(Our) 3 Steps to HAAI





Functional Gradient Boosting

Learn multiple weak models rather than a single complex model



What can be learned?

Natarajan et al (2010, 2011, 2012, 2013) Khot et al (2011, 2013), Yang et al (2016)



What can be learned?

Natarajan et al (2010, 2011, 2012, 2013) Khot et al (2011, 2013), Yang et al (2016), Hadiji et al (2015), Yang et al (2017)



IEEE CVT Dallas, October 15, 2019

Try it yourself

<u>https://starling.utdallas.edu/software/boostsrl/</u>

Tutorial

<u>https://starling.utdallas.edu/software/boostsrl/wiki/</u>

Types of Advice







Types of Advice



Powerful framework that can incorporate different kinds of advice

Odom et al. AAAI '15

Types of Advice

Privileged Information

Odom & Natarajan, Frontiers '18







Knowledge-Based Learning



IEEE CVT Dallas, October 15, 2019 Fung et al. 2002





Active Learning

- Learn initial model from training data m_i
- Generate prediction over data $P_{m_i}(y_i|x_i)$
- Calculate uncertainty $H(P_{m_i}(y_i|x_i))$
- Select example(s) -argmax $H(P_{m_i}(y_i|x_i))$

Odom & Natarajan (2016)

Active Advice Seeking

- Learn initial model from training data m_i
- Generate prediction over data $P_{m_i}(y_i|x_i)$
- Calculate uncertainty $H(P_{m_i}(y_i|x_i))$
- Select example(s) argmax $H(P_{m_i}(y_i|x_i))$

Select clause/rule with the highest uncertainty

Frameworks for Advice Seeking

- Probabilistic Graphical Models
- Relational Probabilistic Models
- Reinforcement Learning
- Inverse Reinforcement Learning
- Imitation Learning
- Probabilistic Planning

Several **<u>Real</u>** Applications



UTSouthwestern Medical Center. VANDERBILT 💱 UNIVERSITY MEDICAL CENTER

The promise of discovery

VV Wake Forest® Baptist Health



Cardiovascular Events Prediction and Treatment



Predicting rare diseases, post-partum depression from survey data



Predicting diabetes / cognition from sensors

Research Institute



Alzheimer's disease prediction





Predicting the side-effects of drugs



Miles to go before we sleep!



- Ensuring Human Trust explain decisions and solicit feedback Always include humans in decision-making
- Enabling Machine Fairness avoid bias in learning (social/economic/religious) impossible to maximize all notions of fairness
- Handling Ethical Issues white lies to make us eat healthy vs negotiation for profit
- Data vs Knowledge what if the evidence is contrary to human perception?
- Optimal/Rational vs. Human-like

Al Serenity Prayer



Human, grant me the serenity to accept the things I cannot learn; Data to learn the things I can; And wisdom to know the difference.

Tweet your questions/comments – @Sriraam_UTD

Thanks to Prof. Rao Khambampati, Arizona State University