Human-Allied Artificial Intelligence

Sriraam Natarajan
The University of Texas at Dallas
This talk is about
This talk is about
This talk is about
This talk is about

Not really!
This talk is about

Facts not fiction 😊
Who we are!
What we do!

- EHR
- Precision Health
- Human-in-the-loop Learning
- Optimization
- Logistics
- Alzheimer's
- Graphical Models
- Reinforcement Learning
- Adverse Drug Events
- Games
- Cardiovascular Health
- Parkinson's Disease Prediction
- Post Partum Depression Prediction
- Logic/Relations
- Information Extraction
- Financial NLP
Can we build systems that can seamlessly interact with, learn from, and collaborate with humans?
Human-Allied AI: The Assistant

Clinical Decision Support System

Primary clinician

- diagnoses, treatment effects, clinical observations

Recommend treatment plans, pull up additional information

Reasoning Module

- EHR
- Home Sensor Data
- Genetic Data

Decision Making Module
Human to machine: Please complete this task!

Example: “Automate physician reports!” or “Enter this data into the electronic health record!”
AI, according to the world: take your data spreadsheet...
...and apply data mining

Gaussian Processes

Latent Dirichlet Allocation

Distillation/LUPI

Graph Mining

Boosting

Autoencoder, Deep Learning

Diffusion Models

Big Data Matrix Factorization

and many more…
Unfortunately, in reality...

Data is messy!
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

IEEE CVT Dallas, October 15, 2019
The Most Important Challenge?

Humans!!!

Humans reason *approximately*
Humans act *unpredictably*

Understanding a human model is *crucial*

*Thanks to Rao Khambampati*
(Our) 3 Steps to HAAI

Learn “only” from data
Effective
Efficient
Generalizable
Personalized
Explainable
... Ignore human knowledge

Allow “richer” human inputs
More than a “mere labeler”
Take advice and guidance
Allows for robust learning

Close the “loop”
Knows what it knows
Asks what it does not know
Student-teacher interaction
Teach the human!

IEEE CVT Dallas, October 15, 2019
Functional Gradient Boosting

Learn multiple weak models rather than a single complex model

\[
\psi_m = \text{Data} - \text{Predictions} + \Delta_m + \text{Induce} + \text{Iterate} + \text{Final Model} = \ldots
\]

- Friedman et al. 2001, Dietterich et al. 2004, Natarajan et al. MLJ 2012

IEEE CVT Dallas, October 15, 2019

What can be learned?


Relational Dependency network

Markov Logic network

Relational CTBN

Learning with Hidden data

Imitation Learning/Relational Policies

Transfer Learning

IEEE CVT Dallas, October 15, 2019
What can be learned?


Multinomial

Poisson

Gaussian

Exponential

Dirichlet
Try it yourself

• https://starling.utdallas.edu/software/boostsrl/

Tutorial

• https://starling.utdallas.edu/software/boostsrl/wiki/
Types of Advice

Monotonicity

As feature $x$ ↑, $P(\text{positive})$ ↑

Precision/Recall Tradeoff

Yang & Natarajan ECML ‘13, Yang et al. ICDM ‘14

IEEE CVT Dallas, October 15, 2019
Types of Advice

Preference Knowledge

Powerful framework that can incorporate different kinds of advice

Odom et al. AAAI ’15

IEEE CVT Dallas, October 15, 2019
Types of Advice

Privileged Information

Training Phase

Deployment/Test

Odom & Natarajan, Frontiers ’18
Knowledge-Based Learning
Knowledge-Based Learning

What advice should the expert give?
passive learning
classical learning setup without any human-in-the-loop guidance
during learning

active learning
learner can query the human-in-the-loop to elicit information
about individual examples, their labels, features

advice-based learning
human-in-the-loop gives general advice including label
& feature preferences, constraints, domain knowledge, rules

active guidance elicitation for learning
human-in-the-loop gives advice about the task including preferences,
constraints, domain knowledge and rules

\[ N_p \]
\[ N_a \approx O(\log N_p) \]
\[ N_{kb} \approx O(\log N_p) \]
\[ N_{ag} \approx O(\log \log N_p) \]
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) - \( \arg\max_{x_i} H(P_{m_i}(y_i|x_i)) \)
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) - $\text{argmax}_{x_i} H(P_{m_i}(y_i|x_i))$

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 **Real** Applications

- Logistics Domains
- Information Extraction
- Games
- Handwriting Recognition
- Image Segmentation and Classification
- Recommendation Systems
- Social Network Analysis

IEEE CVT Dallas, October 15, 2019
Cardiovascular Events Prediction and Treatment

HEART DISEASE is the number 1 cause of death in the U.S., killing 787,000 in 2013

Every 60 seconds someone dies of a cardiovascular disease

In the U.S. someone has a heart attack every 34 seconds

Predicting rare diseases, post-partum depression from survey data

Predicting diabetes / cognition from sensors

Alzheimer's disease prediction

Predicting the side-effects of drugs

Parkinson's disease prediction

IEEE CVT Dallas, October 15, 2019
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**
AI 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