EEP 596: AI and Health Care || Lecture 13

Dr. Karthik Mohan

Univ. of Washington, Seattle

June 1, 2022

Last Lecture

1. Topic modeling and topic segmentation

Today's Lecture

1. Interpretability in Health Care
2. Interpretable models in AI

Explainability

Interpretability and Explainability in health care

1. **Legality**

Interpretability and Explainability in health care

1. **Legality**
2. **Liability**

Interpretability and Explainability in health care

1. **Legality**
2. **Liability**
3. **Confidence in the models**

Interpretability and Explainability in health care

1. **Legality**
2. **Liability**
3. **Confidence in the models**
4. **Wider Adoption**

/ Explainability & Interpretability

Interpretability and Explainability in health care

1. **Legality**
2. **Liability**
3. **Confidence in the models**
4. **Wider Adoption**
5. **Real learning vs spurious learning**

Huskies vs wolves

ICE #1

A/B testing

A clinic wants to incorporate a medical diagnostic tool into its workflow. The idea is to help support doctors making faster diagnosis. The clinic decides to do a A/B test where there is a treatment (using AI diagnostics in pipeline) and a control (status quo). Patients are randomly assigned to treatment or control. After a 4 week study, the clinic saw a 3 % improvement in $F-$score in treatment over control that was statistically significant. Should the clinic go ahead and replace the control with the treatment in their workflow?

a) Yes
b) No
c) Needs further investigation
d) Small difference - so pick either

F-score / Prec ?, Recall?

Explaining Risk to Patients

1. When multiple treatment options, there's a need to explain the pros and cons | Conversation bet doc & patient

Explaining Risk to Patients

1. When multiple treatment options, there's a need to explain the pros and cons
2. Explainable AI helps provide models of data to explain the risks involved with treatments

Explaining Risk to Patients

1. When multiple treatment options, there's a need to explain the pros and cons
2. Explainable AI helps provide data to explain the risks involved with treatments
3. Generic risk vs personalized risk

Explaining Risk to Patients

1. When multiple treatment options, there's a need to explain the pros and cons
2. Explainable AI helps provide data to explain the risks involved with treatments
3. Generic risk vs personalized risk
4. Generic risk: Patients taking this treatment protocol have a 75% chance of success

(Stent procedure) for heart failure

Explaining Risk to Patients

1. When multiple treatment options, there's a need to explain the pros and cons
2. Explainable AI helps provide data to explain the risks involved with treatments
3. Generic risk vs personalized risk
4. Generic risk: Patients taking this treatment protocol have a 75% chance of success
5. Personalized risk: Patients with your kind of profile that take this treatment have a 80% chance of success

Summary statistics

AI can help with

Personalized risk/success model

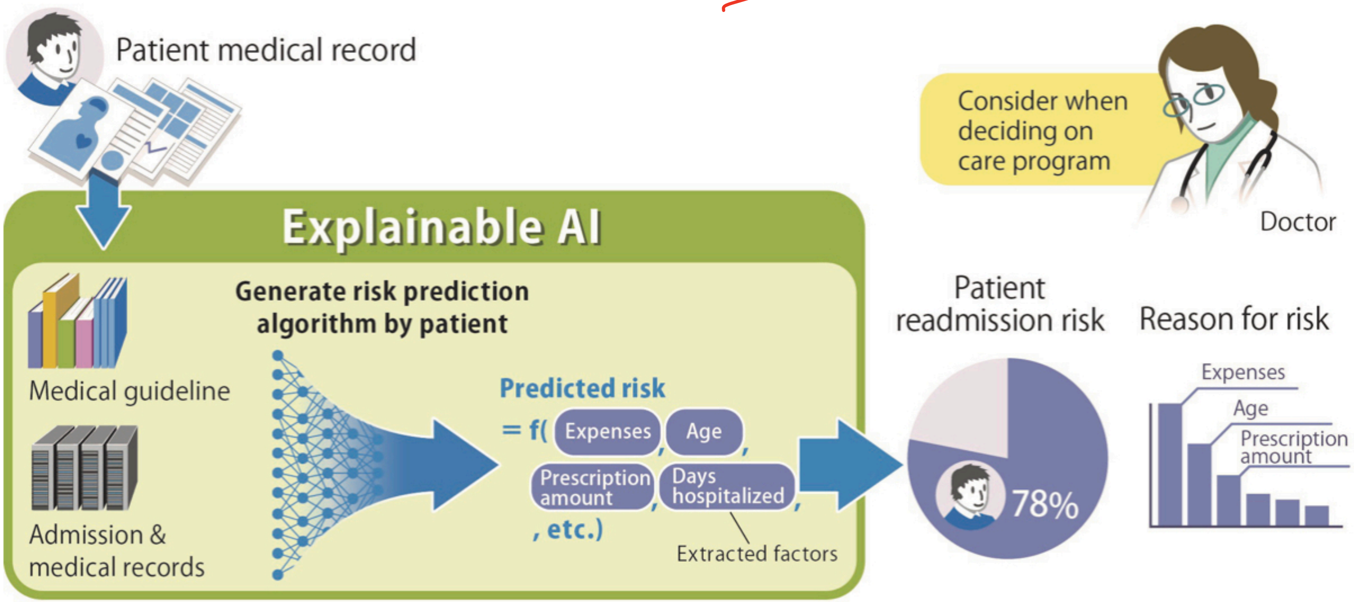
First Line → Chemotherapy
Second Line → Age related

Breakout brainstorm (5 mins)

Let's say you want to build a personalized risk/success assessment tool for different treatment protocols in a cancer clinic. Examples of protocols could include chemotherapy, biological drug therapies, radiation therapy and so on. A patient and their family might aim for a cure for cancer (maximize success) or aim for better quality of life (minimize risk to life). Given that the clinic records patient profiles and their disease and treatment history, how would you go about developing a explainable model to explain the risk/success for different patient profiles? Discuss the data, models and metrics you will use to measure the goodness of your AI explainable model.

Readmission Risk

Cost / time / treatment



Example use situation/case of this AI technology in predicting readmission risk

Readmission Risk Modeling

Saving costs and time

Partners Connected Health and Hitachi announced a AI based tech in 2017 to predict probability of readmission for heart failure patients. This led to significant time and cost savings for patients and hospitals using the model. They use deep models to predict but explainable models to explain it with AUC of 0.7!

2nd model

Good models to explain patient risk?

ICE #2

Which model would more likely be implemented in a clincal setting to explain the personalized risk/success rate of treatments to patients?

a) Deep Learning
b) Logistic Regression
c) Decision Trees
d) k-means

Alternative models to explain black-box AI models

1. Let's say a deep learning model attains high success rate for success assessment of treatment protocols

Alternative models to explain black-box AI models

1. Let's say a deep learning model attains high success rate for success assessment of treatment protocols
2. Clearly, it's not a first choice from explainability standpoint

Alternative models to explain black-box AI models

1. Let's say a deep learning model attains high success rate for success assessment of treatment protocols
2. Clearly, it's not a first choice from explainability standpoint
3. For the sake of explain-ability, can a simpler model be used to correlate input features with the target?

Interpretable vs Explainable AI

Interpretable AI

Refers to models that are inherently interpretable, e.g. small decision trees or linear models with a small number of input variables.

Explainable AI

Refers to the process of applying a method that models the output of a more complex model after training of the complex model.

Explainable AI (XAI) vs Interpretable AI

Simpler model for exaplinability

Breakout brainstorm (5 mins)

You have a deep learning model that gives out personalized success percentage of a treatment protocol for patients. A patient asks to explain why the success of a particular treatment for her is higher than the average success rate? You look towards a simpler ML model to explain the results of a deep learning model. Can you think of designing a simpler ML model that can help identify which input features (e.g. age or medical history, past conditions, etc) are strongly connected to the target (high, medium or low success for example)?

LIME | Local Linear models

DT for explainability

1. Train a DT model on training data along with the black box model

DT for explainability

1. Train a DT model on training data along with the black box model
2. Use the DT model to find a “path” to explain the prediction of the black box model

DT for explainability

1. Train a DT model on training data along with the black box model
2. Use the DT model to find a "path" to explain the prediction of the black box model
3. What are the pros and cons?

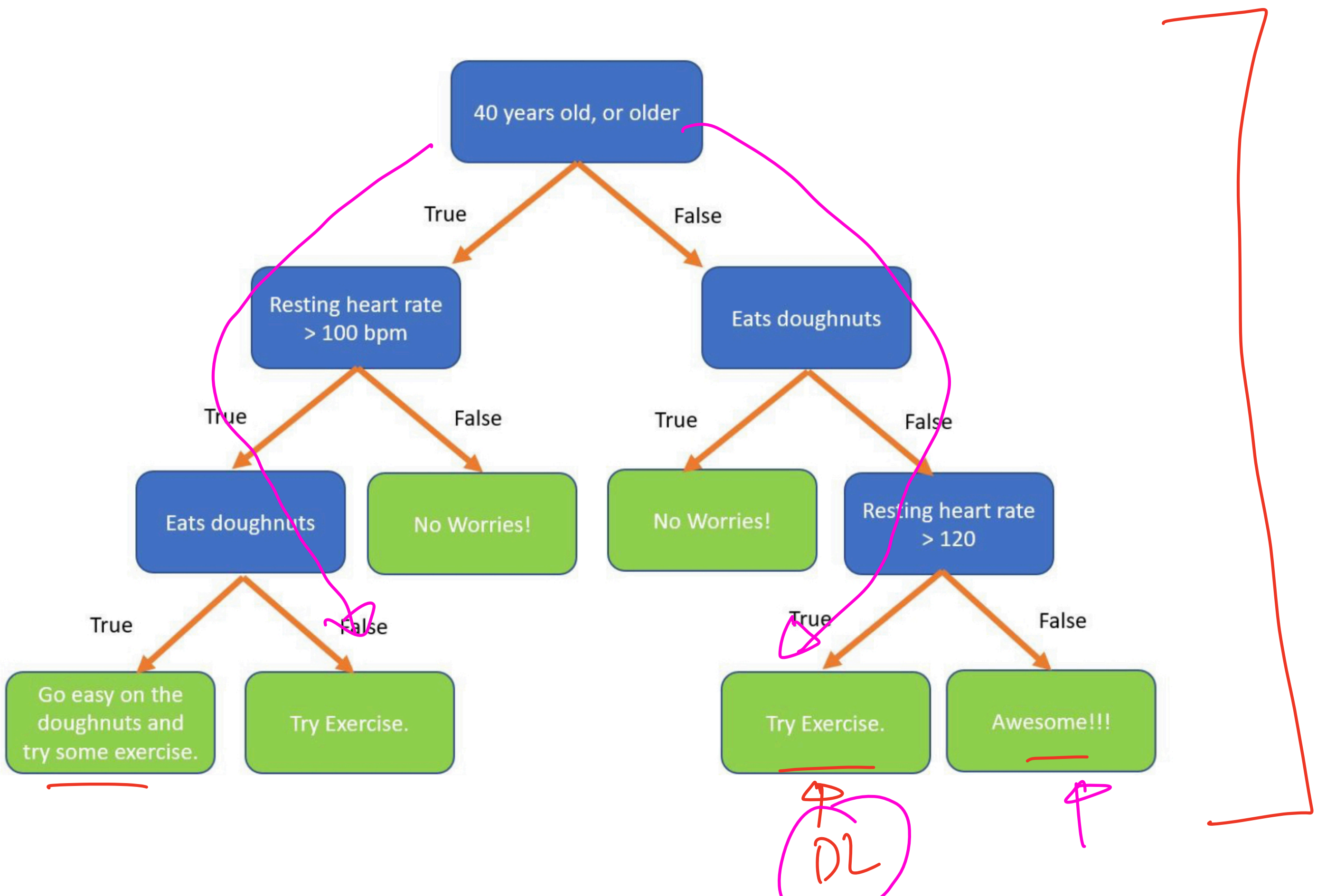
DT for explainability

1. Train a DT model on training data along with the black box model
2. Use the DT model to find a “path” to explain the prediction of the black box model
3. What are the pros and cons?
4. Can we do better?

DT for explainability

1. Train a DT model on training data along with the black box model
2. Use the DT model to find a "path" to explain the prediction of the black box model
3. What are the pros and cons?
4. Can we do better?
5. Find a DT from a random forest that can explain the black-box model the best. More compute time but better explanation!

DT for explainability



Biases in AI

ICE #3

A clinic that primarily sees elderly patients with heart issues is starting to see anomaly in its AI predictions for young patients coming in with heart issues. The 5 year success rate of stent implant in heart for patients having *mytral valve regurgitation* (back flow of blood in heart, which can be fixed by a stent that keeps arteries unclogged) is consistently underestimated in the anomalous predictions. On their training and test data, the AI is able to gain more than 95% accuracy in risk/success assessment. What might be the most plausible reason for the AI model generating anomalous predictions?

a) Random fluctuations in prediction accuracy
b) Not enough data in training
c) Age bias in data
d) Unaccounted factors

Next Lecture

LIME

1. Explaining Deep Learning Models using input attribution in the model
2. Visualization tools to understand deep learning models better

Assignment 3 → June 10!

References

1. Explainability for artificial intelligence in health care: a multidisciplinary perspective
2. Explainable AI model to predict acute illness from EHR
3. Readmission Risk assessment