Machine Learning for Clinicians: Advances for Multimodal Health Data

A Tutorial At MLHC 2018

Michael C. Hughes
Assistant Professor of Computer Science, Tufts University

August 17, 2018

Abstract

This is the accompanying lightly-annotated bibliography to a tutorial at the Machine Learning for Healthcare (MLHC) 2018 conference. Please see the tutorial outline and slides (which this bibliography follows): https://www.michaelchughes.com/mlhc2018_tutorial.html.

Contents

1 Overview 2
2 Making and Evaluating Predictions 5
3 Learning Representations 9
4 Missing Data 15
5 Semi-supervised Prediction 17
6 Multimodal Prediction 19
7 Interpretable Prediction 21
8 Causality 23
9 Reinforcement Learning 24
1 Overview

Tutorials targeted at clinicians:

- “Machine Learning in Medicine” (Deo, 2015) (introduces basic concepts like “supervised” and “unsupervised” learning)
- “Introduction to Machine Learning” (written for audience of Methods in Molecular Biology journal) (Baştanlar & Özuysal, 2014)

Accessible ML Textbooks for Practitioners


Calls to Action:

- “Opportunities for Machine Learning in Healthcare”: (Ghassemi et al., 2018)
- “Machine Learning that Matters” (Wagstaff, 2012)
- “What this Computer Needs is a Physician” (Verghese et al., 2018)

Highlighted recent methods:

- MGP-RNN for Sepsis Risk Prediction (Futoma et al., 2017)

Surveys:

- Survey: Deep Learning for EHR in JAMIA by Xiao et al.
- Survey: “Opportunities and obstacles for deep learning in biology and medicine” by Ching et al. (2018)
- Survey: Deep Learning for Medical Imaging by Litjens et al. (2017)

[Ching et al. 2018] Ching, Travers; Himmelstein, Daniel S.; Beaulieu-Jones, Brett K.; Kalinin, Alexandr A.; Do, Brian T.; Way, Gregory P.; Ferrero, Enrico; Agapow, Paul-Michael; Zeitz, Michael; Hoffman, Michael M.; Xie, Wei; Rosen, Gail L.; Lengenich, Benjamin J.; Israel, Johnny; Lanchantin, Jack; Woloszynek, Stephen; Carpenter, Anne E.; Shrikiumar, Avanti; Xu, Jinbo; Cofer, Evan M.; Lavender, Christopher A.; Turaga, Srinivas C.; Alexandari, Amr M.; Lu, Zhiyong; Harris, David J.; DeCaprio, Dave; Qi, Yanjun; Kundaje, Anshul; Peng, Yifan; Wiley, Laura K.; Segler, Marwin H. S.; Boca, Simina M.; Swamidass, S. J.; Huang, Austin; Gitter, Anthony; Greene, Casey S.: Opportunities and Obstacles for Deep Learning in Biology and Medicine. In: Journal of the Royal Society, Interface 15 (2018), Nr. 141. – ISSN 1742-5662


[Litjens et al. 2017] Litjens, Geert; Kooi, Thijs; Bejnordi, Babak E.; Setio, Arnaud Arindra A.; Ciompi, Francesco; Ghafoorian, Mohsen; Van Der Laak, Jeroen A. W. M.; Van Ginneken, Bram; Sánchez, Clara I.: A Survey on Deep Learning in Medical Image Analysis. In: Medical Image Analysis 42 (2017), p. 60–88. – ISSN 1361-8423


[Xiao et al.] Xiao, Cao; Choi, Edward; Sun, Jimeng: Opportunities and Challenges in Developing Deep Learning Models Using Electronic Health Records Data: A Systematic Review. In: Journal of the American Medical Informatics Association

2 Making and Evaluating Predictions

Evaluating Predictions

Dividing Data into Train/Test/Validation sets and Cross-validation:
- Ch. 7 of (Hastie et al., 2009)
- (Breiman & Spector, 1992)

Evaluating binary classifiers:
- See Zheng (2015)

Intro to ROC analysis: (Fawcett, 2006).

Limitations of Area under ROC curve: (Romero-Brufau et al., 2015) and (Hand, 2009). Also see this blog post by Luke Oakden-Rayner https://lukeoakdenrayner.wordpress.com/2018/01/07/the-philosophical-argument-for-using-roc-curves/

Utility analysis using fixed costs for TP, FP, TN, FN: Blog post by Nicholas Krutch http://blog.mldb.ai/blog/posts/2016/01/ml-meets-economics/

Setting a decision threshold: (Irwin & Irwin, 2011)

Cost curves: (Drummond & Holte, 2006)

Decision curve analysis (Rousson & Zumbrunn, 2011) and (Vickers & Elkin, 2006)

Best practices for model evaluation: (Steyerberg & Vergouwe, 2014)
Making Predictions

Linear Regression: Ch. 3 of (Hastie et al., 2009)

Logistic Regression: Ch 4.4 of (Hastie et al., 2009)

Decision trees : Ch. 2.9 of (Hastie et al., 2009)

Random forests : Ch 15 of (Hastie et al., 2009)

Hyperparameter tuning : See the (unnumbered) chapter of coverage in (Zheng, 2015)

Gaussian processes: (Rasmussen & Williams, 2006)


3 Learning Representations

Bag-of-words Representations

Topic Models

- Topic models survey (Blei, 2012)

Tensor Factorization/Topic Models for EHR

- Marble (Ho et al., 2014b)
- Limestone (Ho et al., 2014a)
- TaGiTeD (Yang et al., 2017)
- PC-sLDA (Hughes et al., 2018)

Learned Image Representations

Convolutional Neural Networks

- Deep CNNs for ImageNet (Krizhevsky et al., 2012)
  - https://www.tensorflow.org/tutorials/images/deep_cnn

Learned Time Series Representations

Highlighted ML+Health Papers

- “Learning to Diagnose with LSTMs” (Lipton et al., 2015)

Hidden Markov Models that do not require aligned time series

- (Liu et al., 2015)
- (Leiva-murillo et al., 2011)
Learned Text Representations

Bidirectional LSTMs:
- (Schuster & Paliwal, 1997)
- (Graves & Schmidhuber, 2005)

1D Convolutional NNs  (Zhang & Wallace, 2015)

Word Embeddings
- GloVe (Pennington et al., 2014)
- word2vec (Mikolov et al., 2013)
- med2vec for EHR codes (Choi et al., 2016)
- Applied to Radiology Report text: (Banerjee et al., 2018)

Tricks of the Trade

Dropout  (Srivastava et al., 2014)

Data Augmentation  Example for Melanoma Classification (Vasconcelos & Vasconcelos, 2017)

Target/Label Replication  Example of LSTM adding loss signal to each timestep, not just final one: (Lipton et al., 2015)

Models that generate data

Denoising Autoencoders
- Denoising AEs (Vincent et al., 2008)
- Deep Patient (Miotto et al., 2016)

Deep generative models and Variational autoencoders:  Johnson et al. (2016) and Kingma & Welling (2014)
GANs: Goodfellow et al. (2014)

medGAN: Choi et al. (2016)


[Graves & Schmidhuber 2005] Graves, Alex; Schmidhuber, Jürgen: Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. In: Neural Networks 18 (2005), Nr. 5, p. 602–610. – ISSN 0893-6080


[Ho et al. 2014b] Ho, Joyce C.; Ghosh, Joydeep; Sun, Jimeng: Marble: High-Throughput Phenotyping from Electronic Health Records via


[Yang et al. 2017] YANG, Kai ; LI, Xiang ; LIU, Haifeng ; MEI, Jing ; XIE, Guotong ; ZHAO, Junfeng ; XIE, Bing ; WANG, Fei: TaGiTeD: Predictive

4 Missing Data

Motivating example:
- Time-of-day of lab tests and 3-year survival rate: (Agniel et al., 2018)

Highlighted Methods:
- MissForest: Random Forest for Imputing Missing data (Stekhoven & Bühlmann, 2012)
- GRU-D: RNNs that handle missingness (Che et al., 2018)
- GAIN: Generative Adversarial Imputation Networks (Yoon et al., 2018)

Other methods
- Generative model that “integrates away” missing data (Caballero Barajas & Akella, 2015)
- (Tresp & Briegel, 1997)


[Che et al. 2018] Che, Zhengping ; Purushotham, Sanjay ; Cho, Kyunghyun ; Sontag, David ; Liu, Yan: Recurrent Neural Networks for Multivariate Time Series with Missing Values. In: Scientific Reports 8 (2018), Nr. 1. – ISSN 2045-2322


5 Semi-supervised Prediction

Methods that can combine few labeled examples with many unlabeled examples.

Evaluation best practices paper

- Realistic Evaluation of SSL (for images) (Oliver et al., 2018)

Highlighted specific methods

- Denoising Autoencoders for 2-stage SSL in EHR (Beaulieu-Jones & Greene, 2016)

Other interesting methods

- Semisupervised with GANs: (McDermott et al., 2018)
- Prediction-constrained training. Longer arXiv version (Hughes et al., 2017)
- Cotraining (Blum & Mitchell, 1998)
- Bayesian co-training and active sensing (given patient demographics data, which one should I image to learn the most): (Yu et al., 2011)


6 Multimodal Prediction

How can we combine images, text, and other modalities of information to develop good learned representations?

Overviews/Surveys on ML methods

- Baltrusaitis, Ahuja, and Morency’s survey: “Multimodal Machine Learning: A Survey and Taxonomy” (focus on images, text, and some video) (Baltrušaitis et al., 2017)


- Another survey: (Ramachandram & Taylor, 2017)

Highlighted ML Methods Papers

- Coordinated embeddings of images and text (vector math with pictures and text) (Kiros et al., 2014)

ML+Health Examples

- Deep Poisson Factor Analysis for Multiple Types of EHR codes (medications, procedures, diagnoses) (Henao et al., 2015)

- MR and PET images for Alzheimer’s: (Lu et al., 2018)

- Cervical cancer images + demographics: (Xu et al., 2016)


[Xu et al. 2016]  XU, Tao ; ZHANG, Han ; HUANG, Xiaolei ; ZHANG, Shaoting ; METAXAS, Dimitris N.: Multimodal Deep Learning for Cervical Dysplasia Diagnosis. In: OURSELIN, Sebastien (Editor); JOSKOWICZ, Leo (Editor); SABUNCU, Mert R. (Editor); UNAL, Gozde (Editor); WELLS, William (Editor): Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016 Volume 9901. Cham : Springer International Publishing, 2016, p. 115–123. – ISBN 978-3-319-46722-1 978-3-319-46723-8
7 Interpretable Prediction

Position papers:

- Doshi-Velez & Kim (2017)
- Lipton (2016)

Classic papers on ML Interpretability in Health:

- (Caruana et al., 2015)

Highlighted Methods:

- SLIM (Ustun & Rudin, 2016)
- LIME (Ribeiro et al., 2016)
- Tree Regularization (Wu et al., 2018)


8 Causality


**Position Paper**
- Pearl on why Supervised Learning isn’t (and won’t be) enough for causal reasoning ([Pearl, 2018](#))

**Tutorials**
- “Causal Inference for Observational Studies” at ICML 2017 [https://cs.nyu.edu/~shalit/tutorial.html](https://cs.nyu.edu/~shalit/tutorial.html)

**Highlighted papers:**
- Counterfactual GP: ([Schulam & Saria, 2017](#))
- Causal-Effect Variational Autoencoder: ([Louizos et al., 2017](#))


9 Reinforcement Learning

Emerging best practices for RL in healthcare are covered in (Gottesman et al., 2018)

Highlighted papers applying RL to real sequential treatment problems in healthcare:

- RL for Sepsis Treatment: (Raghu et al., 2017)
- RL for Schizophrenia: (Shortreed et al., 2011)
- RL for Mechanical Ventilation: (Prasad et al., 2017)

[Gottesman et al. 2018] GOTTESMAN, Omer ; JOHANSSON, Fredrik ; MEIER, Joshua ; DENT, Jack ; LEE, Donghun ; SRINIVASAN, Srivatsan ; ZHANG, Linying ; DING, Yi ; WIHL, David ; PENG, Xuefeng ; YAO, Jiayu ; LAGE, Isaac ; MOSCH, Christopher ; LEHMAN, Li-wei H. ; KOMOROWSKI, Matthieu ; KOMOROWSKI, Matthieu ; FAISAL, Aldo ; CELI, Leo A. ; SONTAG, David ; DOSHI-VELEZ, Finale: Evaluating Reinforcement Learning Algorithms in Observational Health Settings. In: arXiv:1805.12298 [cs, stat] (2018)

