Interpretable Machine Learning for Intensive Care Unit Data

Gareth Booth

Supervised by Sally Shrapnel

Why Intensive Care Unit (ICU) Data?

- Physicians in ICUs have to deal with a massive amount of real time data
- Relevance: Current COVID-19 pandemic
- Want to help overwhelmed ICU doctors to make informed decisions about their patients

COVID-19 Study

- COVID-19 Critical Care Consortium group ECMOCARD study
- ICUs from around the world record and upload data from COVID-19 patients
- Want to know what is important for patients survival
- https://www.uq.edu.au/news/article/2020/05/global-study-of-icu-data-guide-covid-19-treatments

Why Machine Learning?

- State of the art results for predicting ICU outcomes (e.g. death, discharge)
- ICU data is time series data
  - Daily or hourly measurements
  - Lots of machine learning approaches to this problem
- Disadvantage: Need lots of data

Why Interpretable Machine Learning?

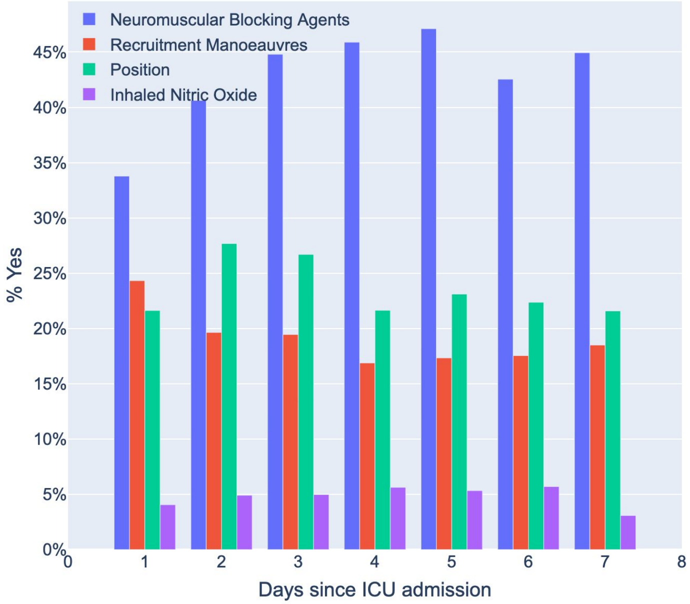
- Machine learning can get a better accuracy than experts in some fields
- But how can we know they are learning sensible relationships?
  - If we train a model to predict death and it uses the 'email' field to make a prediction
- Exploiting biases?
  - Training an image classifier, images only from a certain class show an artifact
- We need a way to see what our model uses to make predictictions (interpretability)
  - Can infer what features are important for this outcome

Expected Outcomes

- Data pipeline for ICU databases (MIMIC-3, REDCap)
- Using Machine learning to predict:
  - Given time series data, predict if death occurred
  - Given admission data, predict time to death, discharge, intubation, sepsis etc
- Explain outcomes
  - What features were most important to make this prediction?

Progress: COVID-19 Data

- Created figures & tables from COVID-19 data
- These will be used in a paper that will be submitted to JAMA



LSTM

- Long Short Term Memory (LSTM) is a type of recurrent neural network
  - Apply all the data in a sequence to the same network
- Feedback connections, both internal state and previous output
- LSTM learns what to pass on as the next state
- Can learn long term dependencies

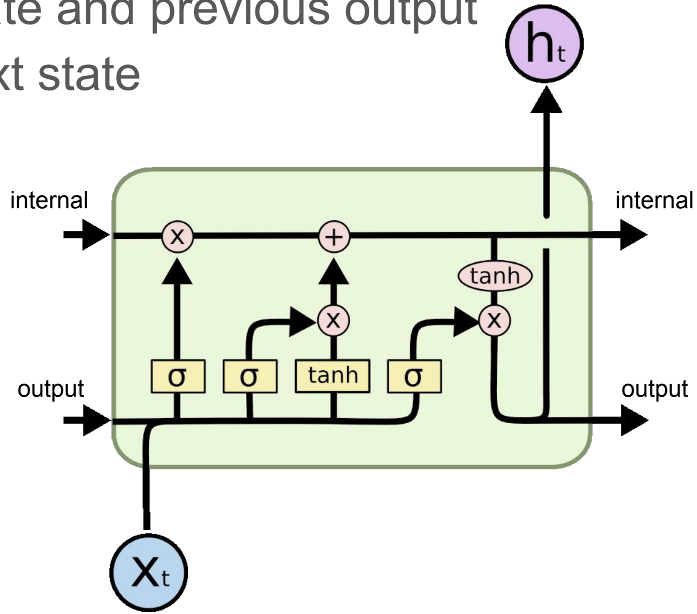


Figure 1: LSTM Unit

MIMIC-III Database

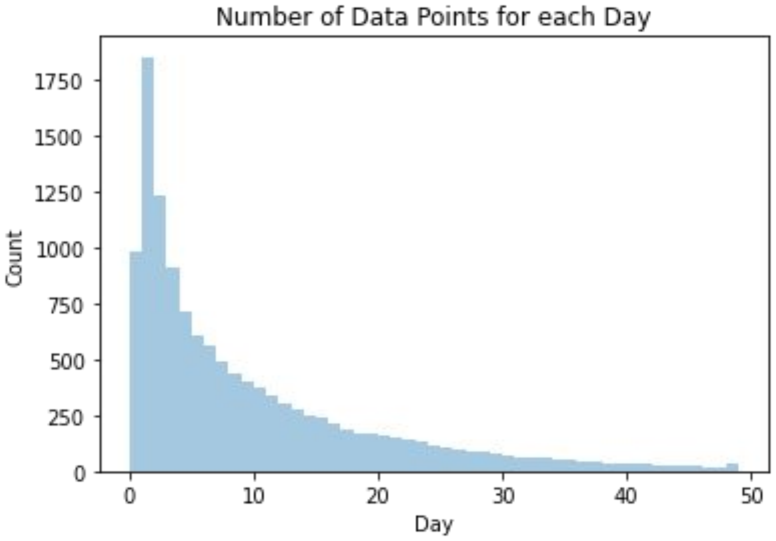
- COVID-19 Critical Care Consortium is an evolving project
    - Not enough data for machine learning (yet)
    - Need to use some other dataset in the meantime
- MIMIC-III
- Open dataset with ~50,000 intensive care unit admissions
- A lot of research has gone into this database
    - Popular for benchmarking machine learning algorithms
- Transfer learning
    - Start with a model trained from MIMIC and use it on COVID-19 data

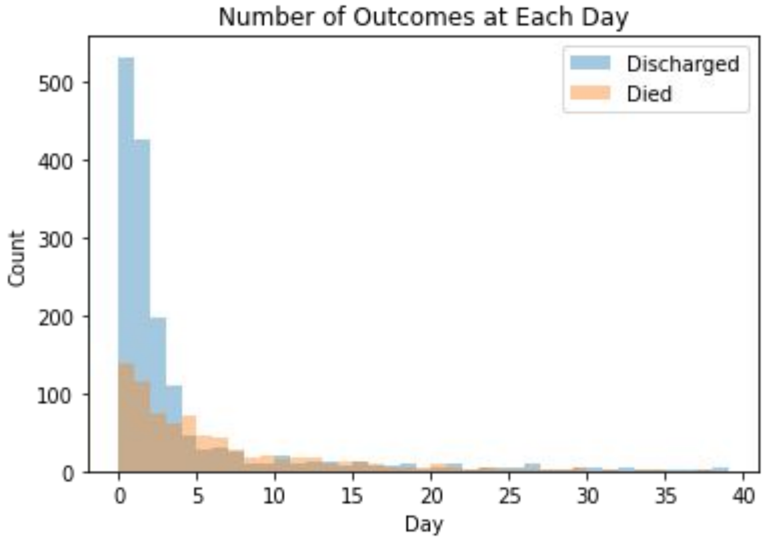
Progress: MIMIC-3 Data

- Prepare data for machine learning
  - Convert from database to csv / dataframe
- Want to try to match the COVID-19 data structure where possible
  - Downsample all features to have daily resolution
- Need to choose relevant features
  - For each patient, drop data from days that have fields missing
  - For now, enough data to drop incomplete data instead of imputing it

Progress: MIMIC-3 Data Training set

- 2368 ICU admissions
- 12819 data points
- Median of 6 days in ICU, average of 10, std of 13
- 30% deaths



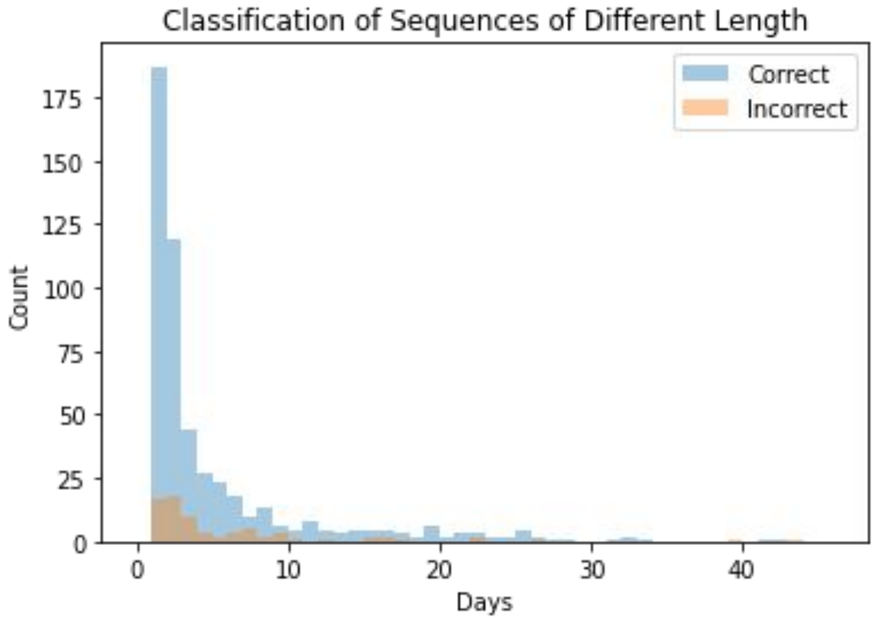


Progress: Predicting Death

- Using the following list of features
  - Creatinine, HCO3, heart rate, haemoglobin, platelet count, potassium, respiratory rate, sodium, age, gender
- Used keras LSTM implementation
- .90 AUC on 20% test set split (by admissions)

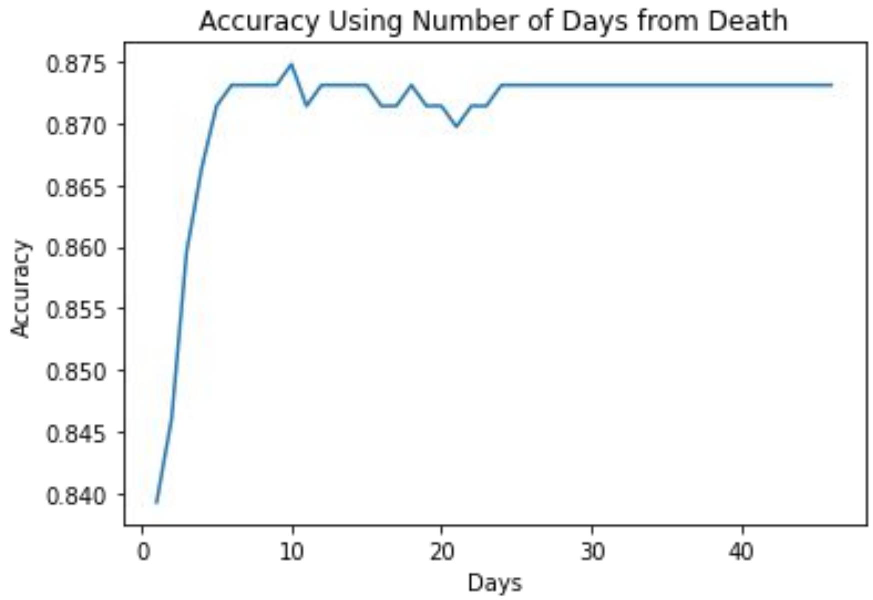
	precision	recall	f1-score	support
Discharge	0.90	0.92	0.91	416
Death	0.80	0.75	0.78	175
accuracy			0.87	591

Progress: Predicting Death



Progress: Predicting Death

- LSTMs can learn long dependencies
- Is the model using more than just the last day to make a prediction?



Limitations

- Daily data
  - Bad resolution, measurement could have been made at a bad time
- Unsure of what the model has learned
  - Is it only looking at the number of days patient has been in ICU for?
  - Is it only looking at the patients age?

Future: Interpretability

- Attention
  - State of the art for LSTM performance, also provides bonus interpretability
  - Created for machine translation, lets the model learn which time steps should attend highly
  - Many different versions of attention exist (Some have been applied to ICUs [1])
  - Will require changing architecture of the model
  - May have to restrict series data to a certain length
  - Attention is noisy [2], so should also use another technique
- SHAP
  - Model agnostic
  - Outputs importance of each feature, taking into account contributions from all features
  - Can be used to provide explanations for a single input sequence
  - Has also been applied successfully in an ICU setting [3]

Future: Other

- Get more features from MIMIC-III
- Use the same model created from earlier on COVID-19 data
    - Hopefully can just plug in the new dataset
- Tackle other prediction problems
    - Admission data is fixed length and has a fixed number of features
        - Random Forests

References

Figure 1: C. Olah, “Understanding lstm networks,” Aug 2015. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[1] D. A. Kaji, J. R. Zech, J. S. Kim, S. K. Cho, N. S. Dangayach, A. B.Costa, and E. K. Oermann, “An attention based deep learning model of clinical events in the intensive care unit.(research article)(report),” PLoS ONE, vol. 14, no. 2, p. e0211057, 2019

[2] S. Serrano and N. Smith, “Is attention interpretable?”arXiv.org, 2019.[Online]. Available: http://search.proquest.com/docview/2238247948/

[3] H.-C. Thorsen-Meyer, A. B. Nielsen, A. P. Nielsen, B. S. Kaas-Hansen,P. Toft, J. Schierbeck, T. Strøm, P. J. Chmura, M. Heimann,L. Dybdahl, L. Spangsege, P. Hulsen, K. Belling, S. Brunak, andA. Perner, “Dynamic and explainable machine learning prediction ofmortality in patients in the intensive care unit: a retrospective study ofhigh-frequency data in electronic patient records,”The Lancet DigitalHealth, vol. 2, no. 4, pp. e179–e191, Apr 2020. [Online]. Available:https://doi.org/10.1016/S2589-7500(20)30018-2

Thanks for Listening

Any Questions?