Deep Learning for Electronic Health Record Prediction

Hao Liu Chunyuan Li Ricardo Henao

Duke University

May 22, 2017

Outline

Data Description

Feature

1. At day0, every people are given injected a virus. And given the electronic health record of several days before day0, we want to know these people's symptom after day0.
2. real value time series data stored in HDF5 format

Data Description

Feature

1. 19 people, every person has a time series data of length about 200000, which is a record of 3 days.
2. time series record contains
   TEMP.csv:temperature in Celsius
   EDA.csv:electrodermal activity in microsiemens
   BVP.csv:blood volume pressure
   ACC.csv:accelerometer,3-axis,acceleration(1/64g)
   IBI.csv:inter beat interval in seconds
   HR.csv:hearth rate from BVP in Hz

Data Description

Label

- 1 x 4 vector for 4 days' shedding titer after day0

mean	[0.22, 1.74, 2.05, 1.63]
variance	[0.91, 2.78, 2.44, 2.17]
median	[-0.5, 1.45, 2.7, 1.7]
min	[-0.5, -0.5, -0.5, -0.5]
max	[2.2, 4.7, 4.7, 4.2]

Data Preprocess

Data Augmentation

- upsampling to augment the data
- normalize the feature
- get 374 data, each with a 20000 x 7 feature, and a 1 x 4 label.

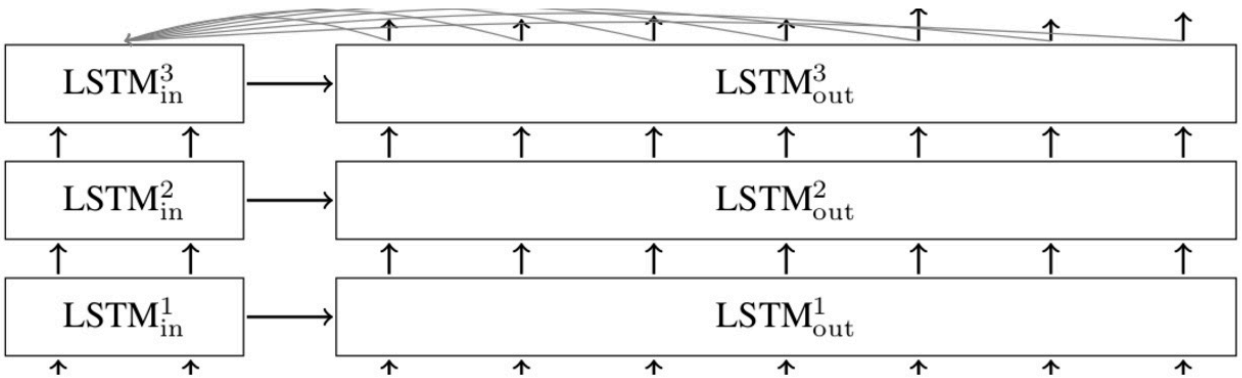
Already Done

- Read the data from the dataset
- Data Preprocess
- Implementing seq2seq, 1D CNN, 1D ResNet and get a preliminary result

Already Done

Model

- Seq2Seq(depth = 3)



- 1D CNN (10 layers)
- 1D ResNet (8 layers)

Already Done

Preliminary Result

- 10 fold cross validation

without Normalization on data feature			
	Loss(MAE)	Correlation	Num of Parameters
Seq2Seq	1.33 ±0.12	0.42 ±0.05	976
1D CNN	1.74 ±0.70	0.26 ±0.34	1175
1D ResNet	1.25 ±0.12	0.42 ±0.06	160652

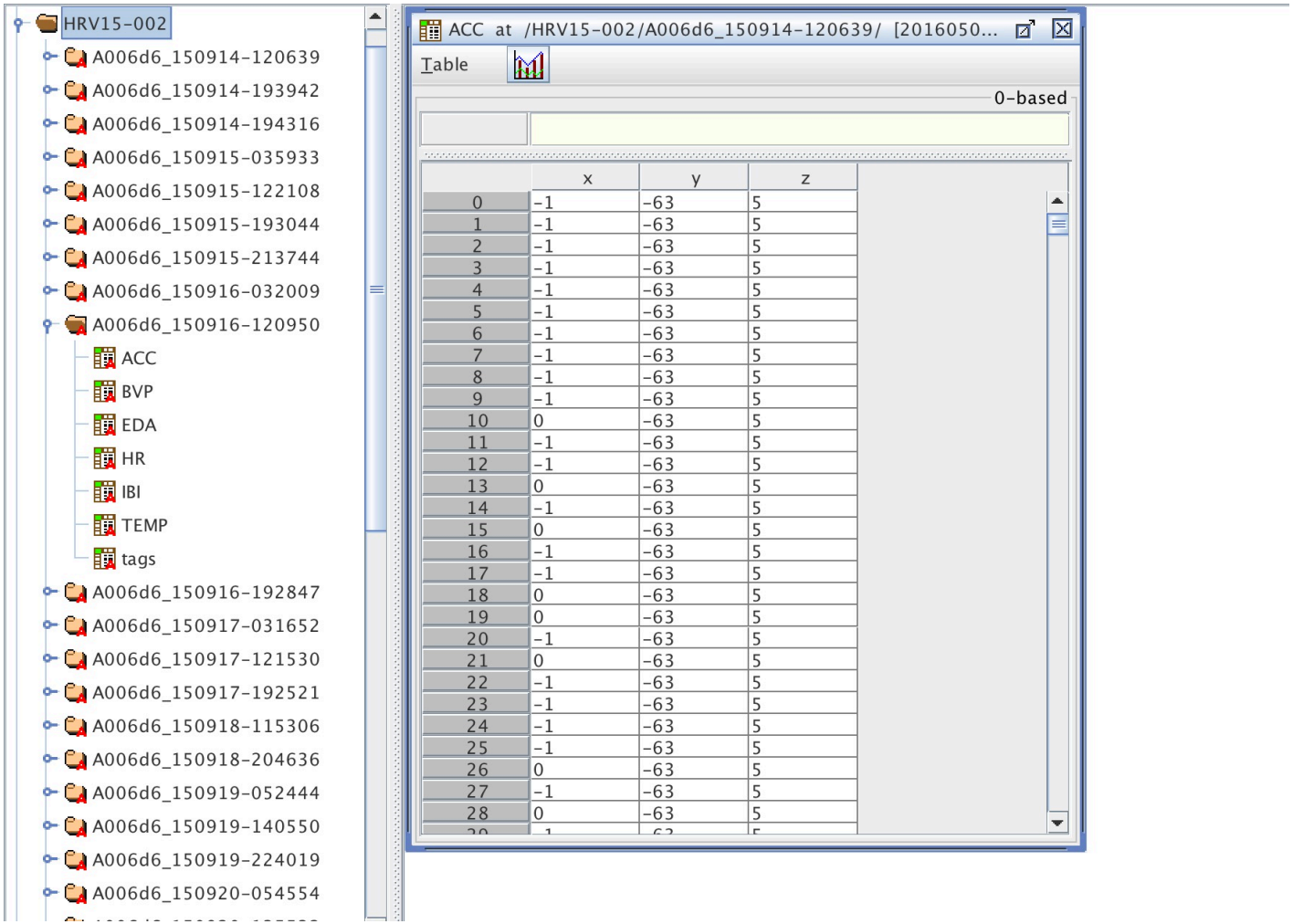
with Normalization on data feature		
	Loss(MAE)	Num of Parameters
Seq2Seq	1.33 ±0.11	976
1D CNN	1.28 ±0.16	1175
1D ResNet	1.60 ±0.53	160652

Future work

- view it as a semi supervised problem, use the sequence before the labels' time as the feature
- better ways for data preprocessing
- other ways for data augmentation
- try some other kinds of networks, such as wavenet
- try attention

Data Description

Feature



	x	y	z
0	-1	-63	5
1	-1	-63	5
2	-1	-63	5
3	-1	-63	5
4	-1	-63	5
5	-1	-63	5
6	-1	-63	5
7	-1	-63	5
8	-1	-63	5
9	-1	-63	5
10	0	-63	5
11	-1	-63	5
12	-1	-63	5
13	0	-63	5
14	-1	-63	5
15	0	-63	5
16	-1	-63	5
17	-1	-63	5
18	0	-63	5
19	0	-63	5
20	-1	-63	5
21	0	-63	5
22	-1	-63	5
23	-1	-63	5
24	-1	-63	5
25	-1	-63	5
26	0	-63	5
27	-1	-63	5
28	0	-63	5

Data Description

Label

subject_id	studyDate	studyDay	sheddingTiter
HRV15-002	2015/9/18	Day1	1.7
HRV15-002	2015/9/19	Day2	3.95
HRV15-002	2015/9/20	Day3	4.45
HRV15-002	2015/9/21	Day4	3.7
HRV15-003	2015/9/18	Day1	0
HRV15-003	2015/9/19	Day2	0
HRV15-003	2015/9/20	Day3	0
HRV15-003	2015/9/21	Day4	0
HRV15-004	2015/9/18	Day1	0.7
HRV15-004	2015/9/19	Day2	1.45
HRV15-004	2015/9/20	Day3	2.7
HRV15-004	2015/9/21	Day4	1.7
HRV15-005	2015/9/18	Day1	0.7
HRV15-005	2015/9/19	Day2	1.7
HRV15-005	2015/9/20	Day3	2.7