Bisht, Suraj

From: b.bardak <b.bardak@etu.edu.tr>
Sent: Sunday, May 1, 2022 3:42 PM
To: Bisht, Suraj
Subject: RE: Multimodal paper: Request for computation details

Hi Suraj,

It is very strange that you take the best results on for example just GRU because I am sure that the proposed CNN is better than using only GRU which I have already reported in the paper.

Also, there are other works that is similar to us which show that using clinical notes or medical entities are works fine rather than just using time-series features:

- Using Clinical Notes with Time Series Data for ICU Management (https://arxiv.org/pdf/1909.09702.pdf)
  - Similar work with us but directly use clinical notes and show that using clinical notes with time-series features is better than just using time series.
- Improving Hospital Mortality Prediction with Medical Named Entities and Multimodal Learning (https://arxiv.org/pdf/1811.12276.pdf)
  - Another similar work with us uses medical entities with Doc2Vec and shows the effectiveness of medical entities.
- Prediction of Mortality and Length of Stay with Deep Learning (https://ieeexplore.ieee.org/abstract/document/9477707/) - This work is written in Turkish because of the language of conference.
  - This study is another study done by us and represents clinical notes directly with Sentence-BERT (used ClinicalBERT). Again using clinical notes are better than just using a single GRU.

Due to the randomness and different seed, there can be difference in the final results however the results that you take with using only GRU is better than multimodal is strange.

I wonder what is the train/val/test size of your experiments?

I am not sure if it will be before your project deadline or not but when I find a free time, I will run my all experiments again and share the patients IDs, and the model weights.

Have a nice day!

Batuhan

2022-05-01 23:19, Bisht, Suraj yazmış:

Hi Batuhan,

We have to submit our final project report to university by end of next week (8th May).

Our executions were performed multiple times with 10 iterations and both with 128 and 256 hidden units.

We have used all_hourly_data.h5 available at GCP.

Due to time limitation we will publish our results and highlight that performance of Base timeseries model and baseline with multimodal are generally higher. Overall performance figures across all the models are higher than published in the paper.

Strange result for the specific experiment for LOS > 5 task where proposed CNN model with Word2Vec embedding is preforming better than best baseline.

Thanks,
Suraj

From: b.bardak <b.bardak@etu.edu.tr>
Sent: Sunday, May 1, 2022 7:30 AM
To: Bisht, Suraj <surajb2@illinois.edu>
Subject: RE: Multimodal paper: Request for computation details

Hi Suraj,

The scores can vary based on seeds which sometimes creates this kind of issue. Due to this reason, I remember that I ran experiments with different seeds and take the mean scores and compare them. However, when I investigate the file that you send, as you mentioned my scores are generally lower than yours but I think that not only for time series but also generally many other settings.

- LOS>3 your averaged word2vec scores are 70.35, 64.69, 56.42 while my reported results are lower than yours 68.63, 61.81, 54.19
- LOS>7 your proposed word2vec scores are 72.45, 20.95, 2.47 while my reported results are lower than yours again with 72.55, 18.78, 1.58

My first question is, how many times did you run these experiments with different seeds? Second, maybe there can be a difference in applying NER operation. Sometimes there can be duplicate NER entities for patients and taking max. 64 entities for patients can be make difference.

There are many different parameters that can effects so it is hard to say directly the reason for this effect. I check the pickles that I stored the results and calculated again and see no difference between my paper and the experimental results. So after you get me back and this issue continues on your side, I can re-run all experiments in my free time and share the model weights with you.

My final question is, do you have any deadline for completing this process? (is this a project or homework? or just to try a new methodology and try to use this framework as a baseline?)

Have a nice day,

Batuhan

2022-05-01 05:25, Bisht, Suraj yazmış:

Hi Batuhan,

Thanks for the quick response, just to let you know that we ran all the experiments (excluding Doc2Vec), surprising most of the baseline Timeseries GRU and with multimodal are performing better than proposed CNN model. One of the observation is GRU baseline performance are way higher than reported in original paper and although proposed model performance stats are in line with paper. Kindly check attached report and share your feedback on what may be reasons for these variation.

Thanks,
Suraj

From: b.bardak <b.bardak@etu.edu.tr>
Sent: Monday, April 25, 2022 6:23 PM
To: Bisht, Suraj <surajb2@illinois.edu>
Subject: RE: Multimodal paper: Request for computation details

Hi Suraj,

This model has been experimented with in my lab server and right now I do not have remote access to that server. I can check it next month when I go to the lab physically.

In addition to this, training your own doc2Vec with clinical note entities is a short process. After extracting entities from patient notes, you assume that each patient's entities like document and train doc2Vec (figure 1). After training doc2Vec, you can get representation for each patient and concat with lstm feature map.

If you have more questions, I can happily help you,

Batuhan

2022-04-26 00:50, Bisht, Suraj yazmış:

Hi Batuhan,

Our team is able to run proposed mode hence kindly ignore below issue.

Doc2Vec is the only open item, do you have ready-to-use doc2Vec model? In repo kindly highlight which code section should be updated to get doc2Vec prediction values?

Thanks,

Suraj

From: Bisht, Suraj
Sent: Sunday, April 24, 2022 11:40 AM
To: b.bardak <b.bardak@etu.edu.tr>
Subject: RE: Multimodal paper: Request for computation details

Hi Batuhan,

Thanks a lot for the response.

Do you have ready to use doc2Vec model? In repo can you highlight which code section should be updated to get Doc2Vec values?

Our team is getting following error in proposed model. Kindly find attached proposed model file we are working on.

```
TypeError                                 Traceback (most recent call last)

<ipython-input-8-fbe1928d4837> in <module>()

     85             #model = textCNN(sequence_model, sequence_hidden_unit,
embed_name, ner_representation_limit)

     86             model = proposedmodel(sequence_model, sequence_hidden_unit,

---> 87                                   embed_name,
ner_representation_limit,filter_number)

     88             model.fit([x_train_lstm, x_train_ner], y_train[each_problem],
epochs=num_epoch, verbose=1,

     89                       validation_data=([x_dev_lstm, x_dev_ner],
y_dev[each_problem]), callbacks=callbacks, batch_size=batch_size)



<ipython-input-7-c79650817411> in proposedmodel(layer_name, number_of_unit,
embedding_name, ner_limit, num_filter)
```

```
     99
    100     #concatenated = keras.layers.Concatenate()([x, text_embeddings])
--> 101     concatenated = merge([x, text_embeddings], mode='concat', concat_axis=1)
    102
    103     concatenated = Dense(512, activation='relu')(concatenated)


TypeError: 'module' object is not callable
```

Thanks,
Suraj

From: b.bardak <b.bardak@etu.edu.tr>
Sent: Saturday, April 23, 2022 12:43 PM
To: Bisht, Suraj <surajb2@illinois.edu>
Subject: RE: Multimodal paper: Request for computation details

Hi Suraj,

Sorry for my late answer. I do not compute the overall computation time but for each iteration, and single-task it is like ~10 min if I recognize it correctly. This means that for 4 separate tasks (in ICU mortality, in-hospital mortality, Los > 3, and Los >7) 10 * 4 = 40 min. And I run the experiments on 10 different seeds which make 40 * 10 = 400 min. But of course, if you want to make hyperparameter tunning or change anything in architecture like word embedding, you need to do all experiments.

For other questions,

* The code is a little bit old, so I do not remember the correct version however, it should be below TensorFlow 2 and should be python 3.6.

* For train doc2vec, I combined all entities of patients which I extracted from notes, and trained doc2vec from gensim myself.

* No, I directly use the original code that I used which should not be got any error. How many times are you getting that error? Maybe it is because of one of the entities that are extracted from NER.

* You need to change the input dimension of the model based on the used embedding method. Other than that, there shouldn't be any errors. If you give some more specific details I can try to help you.

Have a nice weekend,

Batuhan

2022-04-20 07:56, Bisht, Suraj yazmış:

Hi Bardak,

We were able to execute code provided with original paper with some minor updates. Most of the baseline model works fine and need few clarifications

- What are python, keras and tensorflow versions were used for successful code execution?
- For baseline with multimodal approach, Doc2Vec results are clear and there are no references in the code. Kindly guide us how to get results for Doc2Vec embedding?
- Concat embedding is failing, any specific update required on original code? getting following error on concat embedding

  "KeyError: 'all ngrams for word 6u absent from model'"

- Proposed model execution is also pending with few errors with model concatenation/merge, kindly share feedback if you have encounter similar issues with the code uploaded on github

https://github.com/tanlab/ConvolutionMedicalNer

Thanks,
Suraj

From: Bisht, Suraj
Sent: Tuesday, April 5, 2022 11:54 PM
To: 'b.bardak@etu.edu.tr' <b.bardak@etu.edu.tr>
Cc: Bhaskar, Shubhendu <sb59@illinois.edu>
Subject: Multimodal paper: Request for computation details

Hi Bardak,

We are UIUC MCS student, for "CS 598 Deep Learning For Healthcare" course project we are planning to reproduce "Improving Clinical Outcome Predictions Using Convolution over Medical Entities with Multimodal Learning" paper (https://arxiv.org/pdf/2011.12349.pdf)

This paper impression is very good, before reproducing entire paper and running base and proposed models we would to get some feedback on overall computational requirements. In the paper computation resource details are available as "NVIDIA Tesla K80 GPU with 24GB of VRAM, 378 GB of ram and Intel Xeon E5 2683 processor" but we would like to know overall time and computational cost to achieve results available in the paper.

Thanks,

Suraj