Attention-based Neural Text Segmentation for Clinical Notes

Yale

Katherine Sun, Irene Li and Dragomir Radev

Department of Computer Science, Yale University, New Haven, CT

LILY Lab

Introduction

Clinician notes typically contain multiple sections of information, such as medical history, social history, and medications. Currently, there are few papers on text segmentation for clinical data, and no paper has been published for text segmentation based on the MIMIC-III public data set of clinical records (Johnson et al., 2016). The existing text segmentation models for clinical data only perform binary classification at the document-level (Badjatiya et al., 2018; Koshorek et al., 2018). In this work, we build attention-based neural models LSTM and Seq2Seq to segment MIMIC-III clinician notes, which have been pre-labeled with section tags: Beginning (B), History (H), Exams (E), Courses (C), Admission (A), and Discharge (D).

Materials and Methods

We first run binary classification baselines on the MIMIC-III data using the baseline models (Badjatiya et al., 2018; Koshorek et al., 2018). The baseline LSTM, stacked LSTM, and attention model are document-level models with different encoders (Badjatiya et al., 2018). The supervised model contains a sentence embedding sub-network, followed by a segmentation prediction sub-network that predicts a cut-off probability for each sentence (Koshorek et al., 2018). The drawback with these models is that they only perform binary classification to decide the section boundary, failing to specify the category of each sentence.

In comparison with these baselines, we construct both a single-layer LSTM and a bidirectional LSTM that classify a single sentence as 1 to mark the start of a section or 0 if not. We then build upon these LSTMs to perform multiclass classification for a sentence. Finally, we construct a Seq2Seq model composed of an encoder using a bidirectional LSTM, an attention layer, and a decoder. The Seq2Seq model segments an entire document into sections, labeling a sequence of sentences with their appropriate section tags.

Section Type
Begining(B)
History(H)
Exams(E)
Courses(C)
Admission(A)
Discharge(D)

Figure 1. Six possible section tags for MIMIC-III notes.

Baseline Model	Train Accuracy	Test Accuracy
Baseline LSTM	0.9549	0.8059
Stacked LSTM	0.9561	0.8029
Attention Model	0.9915	0.8034
Supervised Model	Pre-trained	**0.8242**

Table 1. MIMIC-III results with baseline models (Badjatiya et al., 2018; Koshorek et al., 2018): binary classification results.

Model	Train Accuracy	Test Accuracy
LSTM: Binary Classification	0.9914	0.9081
BiLSTM: Binary Classification	0.9910	**0.9117**
LSTM: Section Classification	0.9874	**0.8622**
BiLSTM: Section Classification	0.9735	0.8198
Seq2Seq	0.5648	**0.5629**

Table 2. MIMIC-III results with LSTM models and Seq2Seq: multi-class classification results.

Admission Date: [**1801-3-14**] Discharge Date: [**1801-4-12**] **Service:**	B
HISTORY OF PRESENT ILLNESS: This is a 65-year-old gentleman with a history of hypertension. The patient was referred for an exercise treadmill test which was positive and was referred to [**Hospital1 240**] for cardiac catheterization. **PAST MEDICAL HISTORY:** 1. Hypertension. 2. Atrial fibrillation.	H
ALLERGIES: No known drug allergies.	H
MEDICATIONS ON ADMISSION: 1. Aspirin 80 mg p.o. once per day 2. Cartia 300 mg p.o. once per day	A
HOSPITAL COURSE: The patient underwent cardiac catheterization on [**1801-3-14**]. The patient was transferred to the Intensive Care Unit in stable condition. On postoperative day two, the patient was somewhat disoriented, but after several hours the confusion resolved. On postoperative day three, the patient continued to work with physical therapy. By postoperative day five, the patient had completed physical therapy and was stable for discharge.	C
DISCHARGE DISPOSITION: The patient was to be cleared for discharge on postoperative day seven to leave the hospital.	D

Figure 2. Sample text segmentation of clinical data.

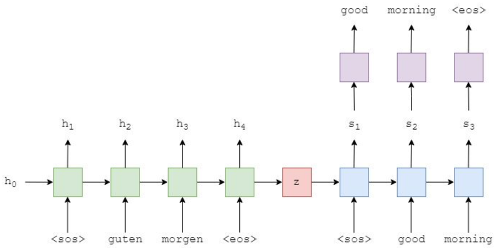

Figure 3. Seq2Seq model architecture.

Results

Table 1 shows the training and test accuracy for each of the document-level, binary classification baseline models. The supervised model from "Text Segmentation as a Supervised Learning Task" performs best, with a test accuracy of 0.8242.

Table 2 shows the training and test accuracy for the sentence-level, binary classification LSTM models; the sentence-level, multiclass classification LSTM models; and the document-level, multiclass classification Seq2Seq model. For binary classification, the bidirectional LSTM slightly outperforms the single-layer LSTM, but for multiclass classification, the single-layer LSTM outperforms the bidirectional LSTM. The Seq2Seq model achieves a test accuracy of 0.5629, lower than the other models but expected because it is both document-level and performs multiclass classification.

Conclusion

In this work, we present several attention-based neural models for text segmentation of MIMIC-III data into sections. We create a set of binary classification LSTM models, a set of multiclass classification LSTM models, and a Seq2Seq model that performs multiclass classification for a sequence of sentences in order to segment a document into sections.

In future work, we will modify the Seq2Seq model to take a sequence of words rather than a sequence of sentences as input.

Acknowledgements

Thanks to Irene Li, Dragomir Radev, and the LILY Lab for their helpful guidance and support.