

Development of Event Detection System with Convolutional Neural Network

Ren Li

E-mail: renli4096@gmail.com



Introduction

With the continuous development of computer and Internet technology, the number of digital data sources and their spreading speed are both greatly increasing. Faced with the "information overload" problem, how to efficiently organize massive information and extract important part is a challenging task. Event as one of the most important components of text, its structure and meaning are no doubt useful to help us understand the text and extract effective information. In this project, we researches Event Extraction (EE) field and develop an event detection system, which aims to automatically detect events described in the text and assign each event an appropriate type.

What is “Event”?

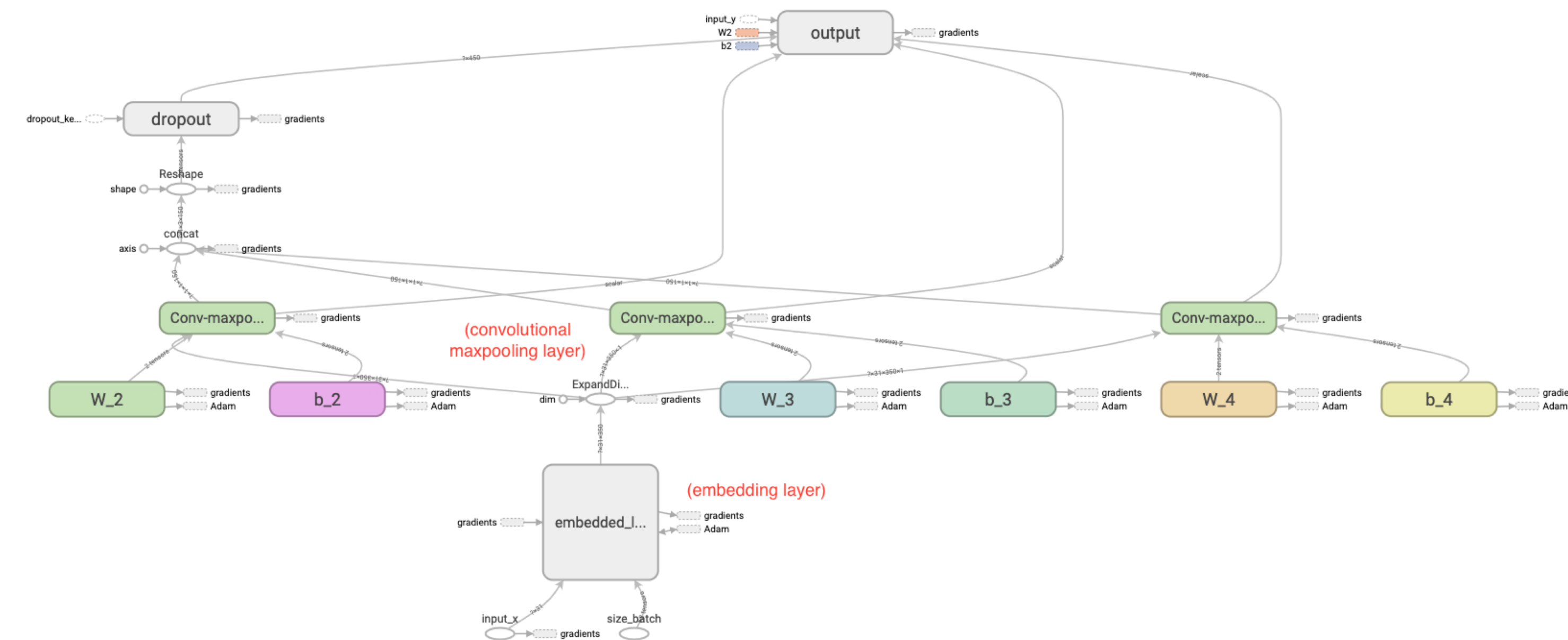
- A specific occurrence, something that happens.
- In text: usually an event can be described by a specific word -- called event trigger / event anchor.
- E.g. “They have been **married** for ten years.” The word married is a event trigger.
- Our aim is to detect the event in the text, i.e. the event word.

Event Types

- Events are classified in 8 types and 33 subtypes in ACE2005 dataset guidelines, which is currently most widely used dataset in EE field.

- ▼ 5. Event Types and Subtypes
 - ▼ 5.1 LIFE
 - 5.1.1 BE-BORN
 - 5.1.2 MARRY
 - 5.1.3 DIVORCE
 - 5.1.4 INJURE
 - 5.1.5 DIE
 - ▼ 5.2 MOVEMENT
 - 5.2.1 TRANSPORT
 - ▼ 5.3 TRANSACTION
 - 5.3.1 TRANSFER-OWNERSHIP
 - 5.3.2 TRANSFER-MONEY
 - ▶ 5.4 BUSINESS
 - ▶ 5.5 CONFLICT
 - ▶ 5.6 CONTACT
 - ▶ 5.7 PERSONELL
 - ▶ 5.8 JUSTICE

System Structure



System Test



(CNN) — The calm of Easter Sunday was shattered by gruesome bombings that killed at least 250 people in Sri Lankan churches and hotels.

Though it's not clear who's behind the eight explosions that forced the country of 21 million people to go on lockdown, they are "certainly acts of terror," said Manisha Gunasekera, high commissioner of Sri Lanka to the UK.

An intelligence memo warning of a possible attack had circulated 10 days earlier, raising questions about whether more preventative measures could have been taken.

"Serious action need to be taken as to why this warning was ignored," Sri Lanka's Minister of Telecommunications, Harin Fernando, tweeted along with a photo of the memo. The document, titled "Information of an alleged plan attack," is dated April 11 and signed by the Deputy Inspector General of Police, Priyalal Dissanayake.

the calm of easter sunday was shattered by gruesome bombings {'Conflict.Attack'} that killed {'Life.Die'} at least 250 people in sri lankan churches and hotels <dot> though it's not clear who's behind the eight explosions that forced the country of 21 million people to go on lockdown <dot> they are <dot> certainly acts of terror <dot> <dot> said manisha gunasekera <dot> high commissioner of sri lanka to the uk <dot> an intelligence memo warning of a possible attack {'Conflict.Attack'} had circulated 10 days earlier <dot> raising questions about whether more preventative measures could have been taken <dot> <dot> serious action need to be taken as to why this warning was ignored <dot> <dot> sri lanka's minister of telecommunications <dot> harin fernando <dot> tweeted along with a photo of the memo <dot> the document <dot> titled <dot> information of an alleged plan attack {'Conflict.Attack'} <dot> <dot> is dated april 11 and signed by the deputy inspector general of police <dot> priyalal dissanayake <dot> the explosions blew out the tiled roofs of churches and hotel windows <dot> killing {'Life.Die'} worshippers and hotel guests in the process <dot> images and footage showed bloodied pews <dot> broken glass <dot> and plumes of smoke <dot> <dot> you can see pieces of flesh thrown all over the walls and on the sanctuary and even outside of the church <dot> <dot> father edmond tillekeratne <dot> social communications director for the archdiocese of colombo <dot> told cnn from st <dot> sebastian's church <dot> one of the explosion {'Conflict.Attack'} sites <dot> he estimated that more than a thousand people had come to the church for easter sunday <dot> because it is a special day <dot> <dot> many came from villages afar <dot> he said <dot>

Figure 1: News about bombing attacks in Sri Lanka from CNN website (will be as the input of system)

Figure 2: System output predictions

Most events are correctly predicted, but there are still some events omitted, like in sentence “people had come to the church for Easter Sunday”, the word “come” is obviously a MOVEMENT event.

Evaluating Indicators

$$Precision = \frac{\text{number of items correctly extracted}}{\text{total number of items extracted by system}}$$

$$Recall = \frac{\text{number of items correctly extracted}}{\text{total number of items need to be extracted by system}}$$

$$F - Measure = \frac{P \times R \times (1 + \beta^2)}{P + R \times \beta^2}$$

F-Measure value can be regarded as a weighted average of Precision value and Recall value.

System Performance

P Value	R Value	F Value
0.7671	0.3237	0.4552

- P Value has already achieved to a good performance.
- R value is low.
- These mean most items predicted are correct, but there are some items the model does not detect out

Conclusions

- The system can work but still needs to be improved.
- System lacks generation ability / adaption ability, i.e. cannot detect out unfamiliar words.
- Solved by increasing amount of training data, trying other neural network structures like Recurrent Neural Network (RNN) or Generative Adversarial Network (GAN).

Applications

In the future, our plan is to revolve around the core algorithm to develop a practical system. The aim of event extraction task is to obtain the events and their participants from text. It will be meaningful to develop an event tracking system that can automatically collect event historical information and track hot event trend. Through the system we can clearly know the development phases of the event without spending much time searching relevant articles, just like a live reporting but you can review it at any time, which will like below:

