**Source code and data description**

**for Syntax-AT-CapsNet**

**1、Code file description**

**（1）Training and testing code files：main\_G\_other、main-G.py**

* main-G.py：Syntax-AT-CapsNet model multi - label text classification data training and testing code files
* main\_G\_other：Syntax-AT-CapsNet model multi - label text classification data training and testing code files

**（2）Model code file**

* network\_G.py：the code of Syntax-AT-CapsNet model and each benchmark model

kimcnn：TextCNN；

capsule\_model\_A：Capsule-A；

capsule\_model\_at：AT-CapsNet；

baseline\_model\_kimcnn\_G：Syntax-CNN；

capsule\_model\_A\_sq：Syntax-CapsNet；

capsule\_model\_at\_sq：Syntax-AT-CapsNet）

* atten\_utils.py：Attention submodule
* GCN\_get.py：Syntactic module
* layer.py：Capsule network layer
* loss.py：Loss calculation

**（3）Data processing code**

* reuters\_process.py：Process multi-label data
* reuters\_process\_other.py：Process single label data
* reuters\_process\_doc.py：Save multi-label text
* reuters\_process\_other\_doc.py：Save single label text
* doc\_to\_adj.py：The text is converted to an adjacency matrix
* adj\_get.py：Construct syntax tree adjacency matrix

**（4）data file**

* Single label data：

subj.hdf5: Subj

subj\_adj.hdf5：Syntax tree adjacency matrix of Subj

rt-polarity.hdf5：MR

rt-polarity\_adj.hdf5：Syntax tree adjacency matrix of MR

custrev.hdf5：CR

custrev\_adj.hdf5：Syntax tree adjacency matrix of CR

* multi-label data

reuters\_multilabel\_all.hdf5： Reuters-Full

reuters\_all\_adj.hdf5：Syntax tree adjacency matrix of Reuters-Full

reuters\_multilabel\_dataset.hdf5：Reuters-Multi

**2、Experimental environment**

* Python: 3.6.9
* pytorch: 1.0.1
* scikit-learn: 0.22
* tensorflow-gpu: 1.5.0
* numpy: 1.18.0
* nltk: 3.4.5
* Keras: 2.1.5
* h5py: 2.9.0
* stanfordnlp: 0.2.0
* stanfordcorenlp: 3.9.1.1

**3、Execution of code**

**（1）Training and Testing Syntax-AT-CapsNet**

* multi-labe：python main-G.py --model\_type capsule-at-sq --dataset reuters\_multilabel\_all
* Single label：python main\_G\_other.py --model\_type capsule-at-sq --dataset subj

**（2）processing data (Modify the path of the corresponding dataset)**

* Process multi-label data：python reuters\_process.py
* Process syntax tree of multi-label data：python reuters\_process\_doc.py; python doc\_to\_adj.py
* Process single label data：python reuters\_process\_ other.py
* Process syntax tree of single label data：python reuters\_process\_other\_doc.py;python doc\_to\_adj.py