

SVAMP

Are NLP Models really able to Solve Simple Math Word Problems?

Dependencies

- compatible with python 3.6
- dependencies can be installed using `SVAMP/code/requirements.txt`

Setup

Install VirtualEnv using the following (optional):

```
$ [sudo] pip install virtualenv
```

Create and activate your virtual environment (optional):

```
$ virtualenv -p python3 venv  
$ source venv/bin/activate
```

Install all the required packages:

at `SVAMP/code:`

```
$ pip install -r requirements.txt
```

To create the relevant directories, run the following command in the corresponding directory of that model:

for eg, at `SVAMP/code/graph2tree:`

```
$ sh setup.sh
```

Then transfer all the data folders to the data subdirectory of that model. For example, copy the ASDiv-A data directory i.e. `cv_asdiv-a` from `SVAMP/data` to `SVAMP/code/graph2tree/data/`.

Models

The current repository includes 5 implementations of Models:

- RNN Seq2Seq at `SVAMP/code/rnn_seq2seq`
 - Basic Encoder-Decoder with Attention Network. Choice of RNN unit provided among LSTM, GRU or RNN.
- Transformer Seq2Seq at `SVAMP/code/transformer_seq2seq`
 - Basic Transformer Network.
- GTS at `SVAMP/code/gts`
 - RNN Encoder with Tree-based Decoder ([Original Implementation](#)).
- Graph2Tree at `SVAMP/code/graph2tree`
 - Graph-based Encoder with Tree-based Decoder ([Original Implementation](#)).
- Constrained Model at `SVAMP/code/constrained`

- Constrained model as described in the paper. Feed-Forward Network maps input embeddings to hidden representations and LSTM Decoder with attention generates the equation.

Datasets

We work with the following datasets:

- `mawps`
 - [Paper](#) and [Github](#).
 - `Data Size:` 2373
 - Evaluated by Cross-Validation over 5 splits.
- `asdiv-a`
 - [Paper](#) and [Github](#).
 - `Data Size:` 1218
 - Evaluated by Cross-Validation over 5 splits.
- `svamp`
 - `Data Size:` 1000
 - Complete challenge set to be used for evaluation.

We have only included the preprocessed ASDiv-A dataset in the `SVAMP/data` directory due to size restrictions. Preprocessed data for all datasets and experiments can be found at the [Github](#) repository for this project.

Usage:

The set of command line arguments available can be seen in the respective `args.py` file. Here, we illustrate running the experiment for cross validation of the ASDiv-A dataset using the Seq2Seq model. Follow the same methodology for running any experiment over any model.

Running Seq2Seq Model for Cross Validation of ASDiv-A

If the folders for the 5 folds are kept as subdirectories inside the directory `../data/cv_asdiv-a:` (for eg, fold0 directory will have `../data/cv_asdiv-a/fold0/train.csv` and `../data/cv_asdiv-a/fold0/dev.csv`),

then, at `SVAMP/code/rnn_seq2seq:`

```
$ python -m src.main -mode train -gpu 0 -embedding roberta -emb_name roberta-base -emb1_size 768 -hidden_size 256 -depth 2 -lr 0.0002 -emb_lr 8e-6 -batch_size 4 -epochs 50 -dataset cv_asdiv-a -full_cv -run_name run_cv_asdiv-a
```

For any clarification, comments, or suggestions please contact [Arkil](#) or [Satwik](#).