

Transformer Analysis on Synthetic Tasks

This repository contains implementations of the Transformer model that we used for Synthetic tasks, i.e. Copy and Counting.

Dependencies

- compatible with python 3.6
- dependencies can be installed using `SAN/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 `SAN:`

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

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

for eg, at `SAN/Transformer:`

```
$ sh setup.sh
```

Models

The current repository includes 2 implementations of Transformer Models:

- Transformer
 - Vanilla Transformer with options to remove residual connections
- DiSAN
 - Directional Self-Attention Network

Usage:

The set of command line arguments available can be seen in the respective `args.py` file.

Running Vanilla Transformer on Copy Task

at `SAN/Transformer:`

```
$ python -m src.main -gpu 0 -max_length 60 -batch_size 128 -epochs 50 -dataset copy_12 -run_name run_copy_task
```

Running Transformer without Encoder-Decoder Residual Connection on Copy Task

at SAN/Transformer:

```
$ python -m src.main -gpu 0 -max_length 60 -batch_size 128 -epochs 50 -no-enc_dec_res -dataset copy_12 -run_name run_copy_task_wo_e-d-res
```

Running Transformer without Decoder-Decoder Residual Connection on Copy Task

at SAN/Transformer:

```
$ python -m src.main -gpu 0 -max_length 60 -batch_size 128 -epochs 50 -no-dec_dec_res -dataset copy_12 -run_name run_copy_task_wo_d-d-res
```

Running Vanilla Transformer on Counting Task

at SAN/Transformer:

```
$ python -m src.main -gpu 0 -max_length 6 -batch_size 32 -epochs 20 -dataset count_data -run_name run_counting_task
```

Running Transformer without Encoder-Decoder Residual Connection on Counting Task

at SAN/Transformer:

```
$ python -m src.main -gpu 0 -max_length 6 -batch_size 32 -epochs 20 -no-enc_dec_res -dataset count_data -run_name run_counting_task_wo_e-d-res
```

Running Transformer without Decoder-Decoder Residual Connection on Counting Task

at SAN/Transformer:

```
$ python -m src.main -gpu 0 -max_length 6 -batch_size 32 -epochs 20 -no-dec_dec_res -dataset count_data -run_name run_counting_task_wo_d-d-res
```

Running SAN on Extended Copy Task

For example the data used is copy_16_20.

at SAN/Transformer:

```
$ python -m src.main -gpu 0 -dataset copy_16_20 -run_name run_san_copy_16_20
```

Running DiSAN on Extended Copy Task

For example the data used is copy_16_20.

at SAN/DiSAN:

```
$ python -m src.main -gpu 0 -dataset copy_16_20 -run_name run_disan_copy_16_20
```