TenGAN

A PyTorch implementation of "TenGAN: Pure Transformer Encoders Make an Efficient Discrete GAN for De Novo Molecular Generation." The paper has been accepted by AISTATS 2024.  Overview of TenGAN

Installation

Execute the following commands:

```
$ conda env create -n tengan_env -f env.yml
$ source activate tengan_env
```

File Description

- **dataset:** contains the training datasets. Each dataset contains only one column of SMILES strings.
    - QM9.csv
    - ZINC.csv
- **res:** all generated datasets, saved models, and experimental results are saved in this folder.
    - save_models: all training results, pre-trained and trained filler and discriminator models are saved in this folder.
    - main.py: definite all hyper-parameters, pretraining of the generator, pretraining of the discriminator, adversarial training of the TenGAN and Ten(W)GAN.
    - mol_metrics.py: definite the vocabulary, tokenization of SMILES strings, and all the objective functions of the chemical properties.
    - data_iter.py: load data for the generator and discriminator.
    - generator.py: definite the generator.
    - discriminator.py: definite the discriminator.
    - rollout.py: definite the Monte Carlo method.
    - utils.py: definite the performance evaluation methods of the generated molecules, such as the validity, uniqueness, novelty, and diversity.

Available Chemical Properties at Present:

```
- solubility
- druglikeness
- synthesizability
```

Experimental Reproduction

- TenGAN on the ZINC dataset with drug-likeness as the optimized property:

```
$ python main.py
```

Citation

```
C. Li and Y. Yamanishi (2024). TenGAN: Pure transformer encoders make an efficient
discrete GAN for de novo molecular generation. AISTATS 2024.
```

BibTeX format:

```
@inproceedings{li2024tengan,
title={TenGAN: Pure Transformer Encoders Make an Efficient Discrete GAN for De
Novo Molecular Generation},
author={Li, Chen and Yamanishi, Yoshihiro},
booktitle={27th International Conference on Artificial Intelligence and Statistics
(AISTATS)},
volume={ 2 3 8 },
year={2024}
}
```