*Study of Deep Generative Adversarial Network Generation of Molecular Graphs*

Anthony Vasquez   
Johns Hopkins University  
Whiting School of EngineeringBaltimor, MD  
[avasque1@jh.edu](mailto:avasque1@jh.edu)

*Abstract*—

Deep Generative Adversarial Networks (DGANs) have demonstrated remarkable effectiveness across various fields, producing high-quality samples that can often deceive both humans and neural networks. For instance, DGANs have been employed in anomaly detection within banking, creating synthetic driving environments for autonomous vehicles, generating synthetic audio and facilitating audio style-transfer, restoring images, and many other applications. One topic of interest, is in the field of chemistry where the challenging problem of generating molecular graphs exists. Many solutions have been tried with varying degrees of success, such as using a variational auto-encoder (Bidisha, 2019). This work compares generations of a variational autoencoder (VAE) to generations of a GAN by training the networks to generate realistic molecular graphs similar to Simplified Molecular Input Line Entry System (SMILES), a database that generates string representations of molecules and reactions (National Library of Medicine, 2024).

Keywords - DGANS, SMILES, Medicine, Chemistry

## Introduction

The generation of new molecules is of interest in medicine, where designing new drugs is long, costly, laborious, and yields few new molecular discoveries (Bidisha, 2019). Some estimates suggest that it takes an average of 10-15 years to develop a single new medicine and comes with a $2.6 dollar price tag (Phrma, 2024). There is much interest in deriving or automating a way to generate physically viable molecules which could potentially speed up the discovery-to-shelf pipeline of pharmaceuticals at a fraction of the price.

A difficulty with working with molecular data is that it is not readily consumable by machine learning algorithms. This is the case with most data, for example Natural Language Processing (NLP) and the use of term frequency-inverse document frequency (TF-IDF), which is a well-known encoding algorithm (Das, 2018). However, there are few vectorization transforms for molecular data. One molecular vectorizer is SMILES, a flexible fragment based molecular representation framework. SMILES consists of codes derived from a breadth-first search on a binary tree consisting of a fragmented molecular graph (Jaun-Ni Wu, 2024). The smiles representation captures the molecular graph which is used to create a representative vector representation as seen in NLP. In this work, a basic GAN will be used to generate valid molecules of a sufficient Novelty, Uniqueness, and Validity as described in Evaluation Metrics.

## Related work

Various methods have been proposed, but two generative methods stand out that explore similar strategies. One method is the VAE, from which many models have been implemented, such as, NeVAE, GraphVAE (Simonosvsky, 2018), GrammarVAE (Kusner, 2017), CVAE (Pagnomi, 2018), SDVAE (Hankun Dai, 2018), JTVAE (Jin, 2019), and CGVAE (Rigoni, 2020), to name a few. This is by no means an exhaustive list, and it can be safely assumed that newer models have been explored. For the sake of brevity, only NeVAE will be discussed, as it is well documented.

The authors of NeVAE claim, “[their] model guarantee[s] a set of valid generated molecules”, and their model can “discover plausible, diverse and novel molecules more effectively than several state-of-the-art methods” (Bidisha, 2019). An innovation used in their experimentations was to use a VAE conjunction with a data representation that captures structural properties of a molecule. Their assumption is that the vector representation embeds N molecular graphs, composed of various sets of atoms (nodes) and bonds (edges), where each may contain various lengths; , where and are nodes and edges respectively, there exists a set of features and edge weighs, , that can be learned (Bidisha, 2019).

Figure 1. Starting from the left, the along with associated features and weights are input to encoder which accumulates information from multiple hops away per node V into an embedding vector to parameterize the posterior distribution based on where the latent representation in the input graph are sampled.

A diagram of a number of squares

Description automatically generated with medium confidence

Their technique creates a continuous latent space, enabling users to extract useful and novel chemical compounds (Bidisha, 2019). Their model is evaluated by decoder generation quality metrics, namely, Novelty, Uniqueness, and Validity. These metrics will be discussed in more depth in Evaluation Metrics.

The other method using a different deep generative paradigm is that of MolGAN, a generative adversarial approach. The GAN architecture was discovered by Ian Goodfellow in 2016, when he described a model consisting of a sample generator and discriminator networks that trained according to an adversarial policy (Goodfellow, 2014). attempts to generate realistic looking samples and attempts to classify if the image came from the distribution of real data or not. At each iteration, the network that loses updates the weights based on the real data distribution. The GAN’s network optimizes according to its respective loss function, and the model is said to converge when can no longer fool . Though GANs were innovative and fast adopted, there was one big problem – mode collapse (Mangalam, 2021). Mode collapse is when the generator fails to produce diverse samples, and it remains a rich topic of research. To combat this, a major improvement to the GAN was proposed with the contributions of WGAN (Arjovsky, 2017). In this paper stability issues were improved on by updating the loss function to the Wasserstein loss, weight clipping, and gradient regularization (Arjovsky, 2017).

MolGAN expands on WGAN by combining reinforcement learning (RL) by rewarding an agent with high objective scores based on several metrics (Nicola De Cao, 2018). Among the objective scores of this method are adversarial loss, to measure the quality of the generated images from the real ones, a loss for molecule solubility, and drug-likeness (Nicola De Cao, 2018). Additionally, an RL component is trained to learn the quality of generated molecules based on “desired chemical properties” (Nicola De Cao, 2018). An objective loss is also calculated for the RL agent during training.

Figure 2. The MolGAN Network has three main components, a generator, discriminator, and an RL component.

A diagram of a graph

Description automatically generated

## ML Architecture

Figure 3. Basic GAN Architecture (Alqahtani, 2019)

A diagram of a generator

Description automatically generated with medium confidence

## Experiment

##### Dataset:

SMILES (Jaun-Ni Wu, 2024)

##### Preprocessing

TBD

##### Experiment

Use a GAN to generate valid molecules. (TBD)

##### Evaluation Metrics

Novelty as shown in Equation 1, is the ability of the network to generate non-existent samples, never seen by the model.

Equation 1. Here, is the set of generated molecules that are valid, and is the train set.

Uniqueness is the measure of unique and valid model generations. Uniqueness is defined as, is describe as the ratio of the set of all generated molecules that are valid by the number of generated molecules. In a sense, Uniqueness attempts to measure how likely a model generated sample is real.

Equation 2. In the Uniqueness measure, is the number of molecule generations for Unique (Bidisha, 2019).

Lastly, Validity is used to measure the degree of the validity of model generated samples. In other words, Validity attempts to measure how likely the generated sample is a real molecule.

Equation 3. Validity measure whether a generation is valid and is .

These metrics are used for both frameworks and are further discussed in the results section.

## Train and Validation Results

TBD

## Out-of-Sample Results

TBD…

## Conclusion

TBD….

# References

Alqahtani, H. E. (2019, Dec 19). Applications of Generative Adversarial Networks (GANS): An Updated Review.

Arjovsky, M. E. (2017, Jan 26). Wasserstein GAN.

Bidisha, S. (2019). NeVAE: A Deep Generative Model for Molecular Graphs.

Das, B. (2018). An Improved Text Sentiment Classification Model Using TF-IDF and Next Word Negation.

Goodfellow, E. A. (2014, June 10). Generative Adversarial Networks.

Hankun Dai, E. A. (2018, Febuary 24). Syntax-Directed Variational Autoencoder for Structured Data.

Jaun-Ni Wu, E. A. (2024). t-SMILES: A Fragment-based Molecular Representation Framework for De Novo Ligand Design. Hunan, China.

Jin, W. E. (2019, March 29). Junction Tree Variational Autoencoder for Molecular Graph Generation.

Kusner, M. E. (2017, March 6). Grammer Variational Autoencoder.

Mangalam, K. G. (2021, Dec 29). Overcoming Mode Collapse with Adaptive Multi Adversarial Training.

National Library of Medicine. (2024, July 22). *National Center for Biotechnology Information*. Retrieved from PubChem: https://pubchem.ncbi.nlm.nih.gov/

Nicola De Cao, T. K. (2018). MolGAN: An Implicit Generative Model for Small Molecular Graphs. Stochholm, Sweden.

Pagnomi, A. E. (2018, December 11). Conditional Variational Autoencoder for Neural Machine Translation.

Phrma. (2024, July 22). *Research & Development Policy Framework*. Retrieved from phrma.org: https://phrma.org/policy-issues/Research-and-Development-Policy-Framework

Rigoni, D. E. (2020, Sep 1). Conditional Constrained Graph Variational Autoencoders for Molecule Design.

Simonosvsky, M. K. (2018, Feb 9). GraphVAE: Towards Generation of Small Graphs Using Variational Autoencoders.