DEEP GENERATIVE MODELS FOR CONDITIONED MOLECULE GENERATION

Members:
Chia Yong Hui (NUS High School of Mathematics and Science)
Esther Toh Jia Ling (Nanyang Girls' High School)
Goh Enxi Xanthe (Raffles Girls' School)

Mentors:
Shen Bingquan, Lim Jing
(DSO National Laboratories)

INTRODUCTION

BACKGROUND

Deep generative models (DGMs) have proved useful in several areas: text, image and music generation.

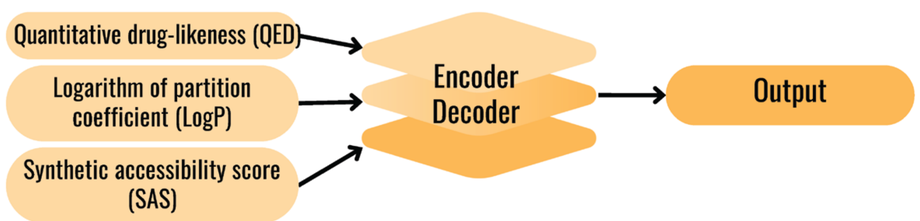
Even in the challenging process of drug discovery!
Potentially saving **25-50%** of time and cost in the discovery and pre-clinical stages.

OBJECTIVE

This project aims to implement a transformer-based architecture with proximal policy optimisation (PPO) and genetic algorithm (GA) to generate drug-like molecules that meet specified molecular properties.

METHODOLOGY

TRANSFORMER



1. Using the ZINC 250k dataset, the properties QED, SAS, and LogP are inputted into the model.
2. The model consists of 6 stacked transformer encoder-decoder blocks. Each block is composed of a multi-head self-attention mechanism and a feed-forward network.
3. The final output is then passed through a linear layer to project it into the vocabulary space. This generates a distribution over possible tokens, from which the next token is sampled.

PROXIMAL POLICY OPTIMIZATION

For each generated molecule, the reward function evaluates how closely its properties match the targets, with the properties assigned different weights:

$$reward_t = 0.45 \cdot reward_{QED} + 0.15 \cdot reward_{SAS} + 0.40 \cdot reward_{LogP} \quad (2)$$

where,

$reward_{QED} = 1 - |QED_{gen} - QED_{target}|$

$reward_{SAS} = 1 - |SAS_{gen} - SAS_{target}|$

$reward_{LogP} = 1 - |LogP_{gen} - LogP_{target}|$

The rewards are used in the calculation of return, which is calculated for each episode and is then used to calculate advantage.
The PPO updates the policy using the clipped surrogate objective:

$$L^{CLIP}(\theta) = \bar{E}_t\,[\min(r_t(\theta) \cdot A_t, clip(r_t(\theta), 1-\epsilon, 1+\epsilon) \cdot A_t] \quad (1)$$

where,

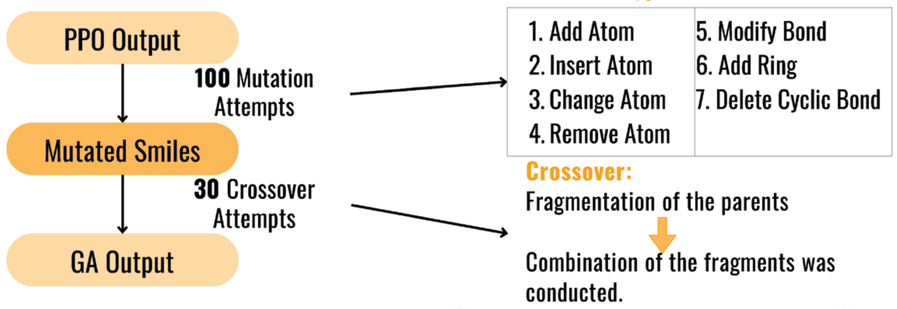
$\bar{E}_t$ denotes the expectation over timesteps

ϵ represents the clipping range that is set to 0.2

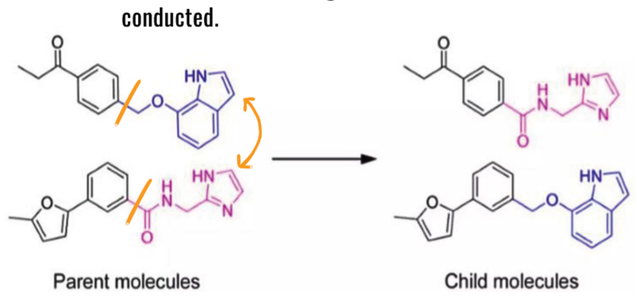
$r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{old}(a_t|s_t)}$ denotes the probability ratio comparing the new policy to the old policy,

where a_t denotes the action taken at time t and s_t denotes the state at time t

GENETIC ALGORITHM



The resultant SMILES are evaluated and saved only if their QED, SAS and LogP values are closer to the target than the original SMILES



RESULTS

EVALUATION

The following metrics were used:

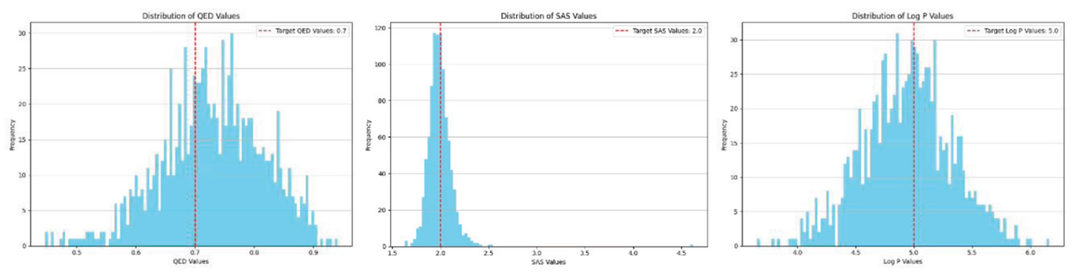
1. Validity: Percentage of valid molecules
2. Diversity: Percentage of unique molecules
3. Novelty: Percentage of molecules that are not in training dataset
4. RMSE (QED,SAS,LogP): Root mean squared error of each property value relative to target value specified

	Validity	Diversity	Novelty	$rmse_{QED}$	$rmse_{SAS}$	$rmse_{LogP}$
Pre-PPO	**97.9%**	98.4%	99.1%	**0.160**	0.139	0.453
Post-PPO	97.6%	**99.1%**	**99.5%**	0.161	**0.136**	**0.403**

Comparison of results for main metrics before and after RL, for target QED: 0.7, SAS: 2.0 and LogP: 5.0

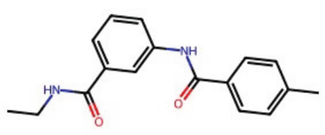
	Validity	Diversity	Novelty	$rmse_{QED}$	$rmse_{SAS}$	$rmse_{LogP}$
Pre-GA	97.6%	99.1%	99.5%	0.161	0.136	0.403
Post-GA	**100.0%**	**99.6%**	**100.0%**	**0.135**	**0.070**	**0.257**

Comparison of results for main metrics before and after GA, for target QED: 0.7, SAS: 2.0 and LogP: 5.0



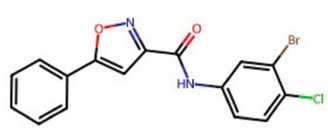
Distribution of QED, SAS and LogP for target QED: 0.7, SAS: 2.0 and LogP: 5.0
Values of the properties are closely centred around the specified targets!

Target: QED: 0.9, SAS: 1.5, LogP: 3



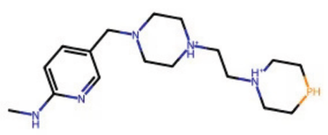
QED: 0.905, SAS: 1.495, LogP: 2.997

Target: QED: 0.7, SAS: 2, LogP: 5



QED: 0.698, SAS: 1.992, LogP: 5.010

Target: QED: 0.55, SAS: 5, LogP: -1.6



QED: 0.548, SAS: 5.027, LogP: -1.599

Top molecules generated for each target after GA

CONCLUSION

- Introduced a hybrid approach, combining a transformer-based model with GA to generate novel drug-like molecules that meet specific molecular properties.
- Demonstrated how RL and BPE can be effective for improving the results for the autoregressive task.
- Future work can explore incorporation of additional biological properties to improve the practical relevance of the generated molecules

ACKNOWLEDGEMENTS

We would like to thank Dr Shen Bingquan and Lim Jing for their guidance and insight.

REFERENCES

D. Sun, W. Gao, H. Hu, and S. Zhou, "Why 90% of clinical drug development fails and how to improve it?," Acta Pharmaceutica Sinica B, vol. 12, no. 7, Feb. 2022, doi: https://doi.org/10.1016/j.apsb.2022.02.002.

"The Drug Discovery Process: What Is It and Its Major Steps," blog.biobide.com. https://blog.biobide.com/the-drug-discovery-process.