MACADAMIA Research Project Report

Human in the loop interaction in reinforcement learning for de novo molecular design

NGUYEN XUAN BINH

Aalto University

MACADAMIA
RESEARCH
PROJECT REPORT

Aalto University
MACADAMIA Research Project Report 2024

Human in the loop interaction in reinforcement learning for de novo molecular design

Nguyen Xuan Binh

MACADAMIA research project submitted in fulfillment of the requirements for the degree of Master of Science.
Otaniemi, 1st May 2024

Advisor: Yasmine Nahal, Prof. Samuel Kaski

Aalto University
School of Engineering
Master's Programme in Science and Technology

Author
Nguyen Xuan Binh

Title
Human in the loop interaction in reinforcement learning for de novo molecular design

School School of Science

Degree programme
Master of Science in Computer, Communication and Information Sciences (CCIS)

Major Machine Learning, Data Science and Artificial Intelligence (MACADAMIA)

Advisor Yasmine Nahal, Prof. Samuel Kaski **Code** SCI3044

Date 1 May 2024 **Pages** 24 **Language** English

Abstract

De novo molecular design, which makes use of computational techniques to engineer novel, diverse, synthesizable and drug-like molecular structures, is an essential step in the drug discovery process. The formulation of human-in-the-loop workflow and comparison of three rival feedback mechanisms — the scoring model, the comparing Bradley-Terry model, and ranking ListNet model, represents the main contribution of this study. The scoring model directly estimates the Dopamine Receptor D2 (DRD2) probability for each molecule, while the Bradley-Terry model evaluates the likelihood that one molecule is better than another in terms of DRD2 probability. Finally, preference scores are assigned to rank sets of molecules by the ranking ListNet model.

We tested these models under two noise levels (0.0 and 0.1) and three different acquisition functions (random, uncertainty, and greedy) to assess how well they perform. According to our research, the ListNet model consistently performs better than the other models in terms of molecular diversity and drug-likeness, the Bradley-Terry model performs best in terms of novelty criterion, while the baseline scoring model excels most in chemical property-targeted molecules generation and synthesizability. This work shows how customized feedback model mechanisms can improve de-novo molecular design and eventually lead to the identification of therapeutic compounds with ideal characteristics.

Keywords Reinforcement Learning, Human in the Loop, De Novo Molecular Design, REINVENT, RDKit, Cheminformatics, SMILES

urn https://aaltodoc.aalto.fi

Contents

1 Introduction

1.1 Introduction to cheminformatics

Cheminformatics is the study of how chemical and biological information is represented and used on computers. This field has numerous applications, including drug discovery, healthcare, data mining, and a variety of other areas. Caffeine, nicotine, adderall, quinine, morphine, benzylpenicillin, tamoxifen, Zantac, Prozac, Valium, THC, and cortisone are all approved and frequently used medications that help to sustain human health.

Researchers frequently represent compounds as molecular graphs, which is extremely useful for data storage and analysis. For example, benzodiazepines, a type of sedative, share a shared ring system or scaffold, indicating that similar compounds frequently have similar bioactivity. Chemical databases, such as the PubChem database of the National Institutes of Health (NIH) Library of Medicine, are industrial-standard resources for drug discovery and other uses. [1].

To represent molecules on computers, researchers usually use these two formats:

- Molecular depiction (2D): it shows the arrangement of atoms and the bonds between them in a planar (flat) format, like in the chemistry textbooks [2].
- Simplified Molecular Input Line Entry System (SMILES): The atoms and bonds are represented in a single line using ASCII string [3]. For example, the SMILES for water (H_2O) are "O". One advantage of SMILES is that it can be easily transformed to 2D or 3D structures using cheminformatics tools like RDKit for visualization. Furthermore, the SMILES method assures that each molecule has a unique representation because it adheres to defined standards for atom ordering and bond traversal.

When searching for molecules, we can look for a list of molecules with the same substructure. In other words, we can apply the SMILES arbitrary target specification (SMARTS), which is analogous to a regular expression. The SMARTS line notation is expressive and enables for highly exact and transparent substructural specification and atom typing [4]. Considering the SMARTS notation $[\#6]12 \sim [\#7] \sim * \sim [\#7] \sim [\#6].1 \sim * \sim * \sim * \sim *2$ represents the imidazole pattern subgraph matching. The asterisk is a wildcard placeholder for any atom, while the tilde symbol indicates that any sort of bond between atoms in a substructure is accepted.

1.2 De Novo drug molecular design

De novo molecular design is the process of creating new molecular structures from scratch using computational approaches rather than depending on pre-existing templates. This method plays a major role in drug development since it allows the synthesis of new molecules that have specific desired features. Advanced algorithms are often used to explore chemical space, producing and optimizing novel molecules based on established criteria such as biological activity, drug-likeness, and synthesizability. For example, Gillet (2000) proposes to employ evolutionary algorithms in de novo molecular design, indicating their utility in pharmacophore mapping and receptor modeling to develop drug-like molecules from enormous chemical spaces [5]. Similarly, Hsu et al. (2019) describe a new data format for computational molecular design, which enables the production of a comprehensive library of new compounds [6]. Therefore, researchers can find molecules with the best theurapeutic qualities by searching the huge chemical space via techniques like evolutionary algorithms and novel data structures.

1.3 Human-in-the-loop De Novo drug design

The 'human-in-the-loop' framework has been rising in popularity because it allows humans to use their domain expertise into the modeling process, connecting computer science, cognitive science, and psychology [7]. Reinforcement learning from human feedback (RLHF) is a popular framework that learns from human feedback instead of a designed reward function. It is a practical alternative that introduces a critical human-in-the-loop component to the standard RL learning paradigm [8]. RLHF varies from RL by allowing humans to define and refine the objective in the loop, rather than specifying it beforehand.

Besides RLHF, there are other diverse HITL methods on de novo molecular design as well. For example, the Chemical space-based de novo design method allows researchers to specify desired features and explore across the huge chemical space to discover interesting molecules [9]. DrugMint, a website that predicts and designs drug-like molecules, uses predictive models to help scientists identify compounds with high drug-likeness scores [10]. DrugChat, a platform that uses graph neural networks and massive language models to provide feedback and ideas, enables users to query and revise chemical designs interactively [11]. These HITL approaches all show great potential for incorporating human expertise into de novo drug discovery when designing objective functions is challenging.

2 Research project

2.1 REINVENT: AI for molecular de novo design

REINVENT is an advanced AI tool designed for de novo drug design, providing a comprehensive platform for generating novel drug-like molecules [12]. The primary motivation behind REINVENT is to efficiently explore vast chemical spaces, identifying and optimizing molecules with desired biological activities and drug-like properties. This is motivated by the fact that traditional drug discovery methods are often time-consuming and costly, which REINVENT aims to address.

Regarding usage, REINVENT allows researchers to define target properties and constraints, and then it generates molecules that meet these criteria. Users can interact with the platform, providing feedback to refine and improve the generated compounds. This human-in-the-loop approach ensures that the generated molecules are novel, diverse and synthetically practical [13].

REINVENT's iterative optimization involves multiple steps as follows [14]:

- Initialization: Users define the initial set of constraints and desired properties to optimize. This may include diversity filter, inception, chemical properties, configurations and most importantly, the scoring function. This is the basis for developing different feedback to REINVENT. All of these settings are written to a json file, which would be read by REINVENT.
- Generation: The AI model generates a batch of candidate molecules at each optimization step. Default value for the batch size is 64 and default value for number of optimization steps is 100.
- Evaluation: Generated molecules are scored based on the scoring function (or the feedback function). REINVENT tries to generate molecules that maximize the scores.
- Optimization: The tool refines the molecules through multiple steps, and finally return SMILES in the scaffold memory csv file.

2.2 Project motivation

In this project, the general objective is to use REINVENT to generate molecules that has dopamine receptor D2 (DRD2). This receptor is crucial to regulating dopamine, a neurotransmitter associated with motor control, motivation, and reward mechanisms [15]. Optimizing presence of DRD2 could provide improvements

to inhibit diseases such as schizophrenia, Parkinson's disease, and drug addiction. Specifically, antagonists of DRD2 appears to reduce tumor growth in cancer, making it a valuable property for developing anticancer treatments [16][17].

Nonetheless, the primary motivation for this project is to generate molecules that not only have high DRD2 probability but are also novel (having new, possibly unseen scaffold structures) and they should also have drug-like properties. This is important for drug discovery to explore new chemical spaces that may unexpectedly have interesting properties. For instance, Wang et al. (2024) emphasize the need for novelty and drug-likeness in molecular design [18], while Fotie et al. (2023) discuss the importance of drug-like properties in drug design [19].

Another motivation of this project is to establish a human-in-the-loop (HITL) workflow, which involves human expertise to guide the molecular design process. This approach allows chemists to provide comparative feedback, which is often easier than direct scoring. Some studies have demonstrated the impact of HITL in enhancing molecular design through comparative feedback mechanisms [20].

The final motivation for this project is to develop a surrogate ML model that can decently predict DRD2 probability for novel molecules. This is not of major importance because large datasets already exist, containing hundreds of thousands of molecules labeled with DRD2 presence and there is already a pretrained model TDC Oracle that can measure DRD2 probability with high accuracy.

2.3 Project problem statement

This research project aims to carry out these experiments

- Use REINVENT software as a reinforcement-learning model for generating molecules in SMILES format that maximizes probability that they have DRD2. In other words, this project actually does not develop any reinforcement learning framework, as it is already an existing model in REINVENT.
- Develop three distinct human feedback mechanism: the scoring, comparing and ranking feedback. Later, we can show that human in this context is not necessarily an actual human, but an ad-hoc surrogate ML model that acts as a human agent. This is the central contribution of this research project.
- Develop a human-in-the-loop workflow that couples with REINVENT to generate novel molecules with preferable scores of chemical properties. This is the second contribution of this project.

- Develop an active learning framework where promising SMILES are selected for human feedback using various acquisition functions, with either perfect labelling or noisy labelling. This is the third contribution of this project.
- Run the HITL workflow, benchmark all running cases and report discoveries.

Unfortunately, due to the time scope, the author cannot afford to build a Graphic User Interface for modelling human feedback.

3 Methodologies

3.1 Human-in-the-loop workflow

This project workflow is heavily influenced by the work of Iiris et al. [21]. In fact, the workflow is identical to the second case study (Human Chemist Component). The main difference is that various feedback mechanism is tested to observe the properties of generated molecules, while in the paper of Iiris, it seems that only the feedback of preference (yes/no) is tested. Another difference is this workflow directly uses the human component to score the molecules, and use an Oracle to label the DRD2 probability, where Oracle's parameters stay fixed throughout the process. The labelling agent in the original workflow, in constrast, gets updated. To recap, this workflow involves an AI that helps a chemist to decide parameters of an multi-parameter optimization function $S_{r,t}(x)$ iteratively at round r and iteration t, where r are rounds of molecule generation with REINVENT, and t are number of active-learning interactions with a human component to receive feedback for generated molecules [21]. The objective consists of K molecular properties $c_k(x)$ with relative weights w_k. The score of the kth property is measured using a scoring function $\phi_{r,t,k}$. Since we only maximize DRD2, we would have $k = 1$ here.

In reality, the humans usually only have a vague intuition of DRD2 presence initially, and the humans would become more proficient over time when REINVENT recommends better SMILES that have DRD2. Furthermore, because evaluating or comparing thousands of molecules is extremely time-consuming even for experienced chemists, we are obliged to use an ML model to act as a surrogate human. To ensure that the ML imitates a human, it should classify DRD2 with very low accuracy (around 0.5) at the beginning of the process. Gradually, during interactions, the ML model would acquire new molecules to be labelled, based on the scheme of urgency (random, uncertainty or greedy). Below is the pseudo-code from the paper but was directly modified to fit in this research work [21].

Algorithm 1 Human-in-the-loop workflow

Require: An Oracle that reliably estimates DRD2 probability, a weak ML model that returns feedback as a scoring component $S_{\theta_{0,T}}$, number of REINVENT rounds R, number of human interactions T, number of queries Q at each interaction, acquisition function ACQ, Oracle's noise level σ

$D_{0,T} \leftarrow \emptyset$ ▷ Initially, the training dataset is empty
for $r = 1, 2, \ldots, R$ **do** ▷ Looping over REINVENT rounds
  $S_{\theta_{r,1}} \leftarrow S_{\theta_{r-1,T}}$ ▷ Current round ML model is the ML model from last interaction of previous round
  $D_{r,1} \leftarrow D_{r-1,T}$ ▷ Current training dataset is the dataset from last interaction of previous round
  $U_r \leftarrow$ REINVENT$(S_{\theta_{r,1}})$ ▷ U_r: set of molecules from REINVENT using the ML model $S_{\theta_{r,1}}$
  $U_{r_{\text{best}}} \leftarrow$ Select top n_{best} molecules x with highest scores from U_r
  for $t = 1, 2, \ldots, T$ **do** ▷ Looping over online interations with ML model
    for query $= 1, 2, \ldots, Q$ **do**
      $x^* \leftarrow$ ACQ$(S_{\theta_{r,t}}, U_{r_{\text{best}}})$ ▷ Obtain new SMILES using the chosen acquisition function ACQ
      $y^* \leftarrow Oracle(x^*) + \mathcal{N}(0, \sigma)$ ▷ Acquire feedback y^* of DRD2 probability for x^* SMILES from Oracle plus some noise
      $U_{r_{\text{best}}} \leftarrow U_{r_{\text{best}}} \setminus x^*$ ▷ Remove x^* SMILES from $U_{r_{\text{best}}}$
      $D_{r,t} \leftarrow D_{r,t} \cup \{(x^*, y^*)\}$ ▷ Update the dataset with new queries
    end for
    $S_{\theta_{r,t}} \leftarrow S_{\theta_{r,t}}$ retraining on $D_{r,t}$ ▷ The ML model is updated
  end for
end for

A question may arise from this workflow is that, if there is already access to the Oracle that can reliably tells DRD2 probability, why should one invest time to train a weak model from scratch, when we can directly use the Oracle as the scoring component in REINVENT. The motivation for this workflow is that we assume the function $Oracle(x^*)$ is an expensive, a blackbox function or an erroneous human chemist, and we are trying to build an ML model that is cheaper to infer, easier to interpret or more robust against noise. To conclude, the settings used in this project is $R = 3$ REINVENT rounds, $T = 5$ HITL interactions, $Q = 56$ queries, REINVENT's batch size of 64 with 100 optimization steps. Initial training dataset size is 200 molecules of balanced class (100 with and without DRD2), and after 3 REINVENT rounds of 5 iterations, the final training dataset size becomes $200 + 3 \times 5 \times 56 = 1040$ SMILES for the last ML model.

3.2 Feedback model architecture

This project develops three feedback models for REINVENT during the scoring stage: the scoring (baseline) model, the comparing (Bradley-Terry) model and the ranking (ListNet) model. While they differ in how they return the feedback to

REINVENT, they has essentially the same neural network architecture, which is to learn the DRD2 probability. This aspect is convenient because we would prefer these feedback models to not only offer diverse feedback to REINVENT but they should also be used to infer DRD2 probability of new SMILES quickly. The figure below shows the simple architecture for learning DRD2 probability, where the Sigmoid activation ensures that the logit is restricted to range [0, 1].

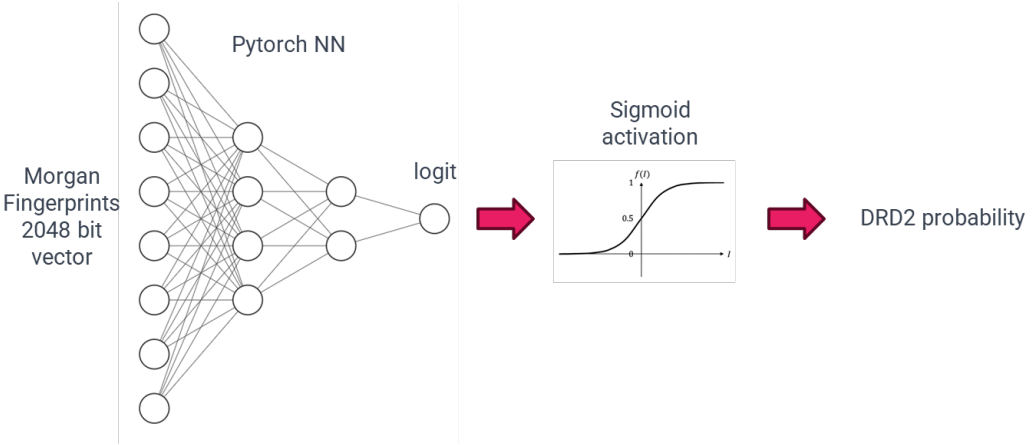

Figure 1. *ML model architecture for learning DRD2 probability*

As a result, the feedback model has three distinct types of outputs: DRD2 probability output, DRD2 individual feedback, and DRD2 batch feedback. DRD2 probability output is only used when we need to classify new SMILES whether they have DRD2. Details of the three feedback models are demonstrated below

Scoring model

The scoring model predicts the probability that a given SMILES string has DRD2 activity based on its Morgan Extended-Connectivity Fingerprint (ECFP). This is a straightforward feedback that REINVENT usually expects from users.

Inputs: Morgan fingerprints of 1 single SMILES.
Outputs: Probability that the SMILES have DRD2 (score between 0 and 1)
Individual feedback: The same as outputs
Batch feedback for REINVENT and acquisition evaluation: The same as outputs for all molecules.
Loss function to train Pytorch model: BCELoss, which has formula
$\text{BCE}(p, q) = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(p(y_i)) + (1 - y_i)\log(1 - p(y_i))\right]$

Comparing model, using Bradley-Terry formulation

As mentioned earlier, the motivation for comparing molecules stem from the idea that humans may have difficulty in telling the exact DRD2 probability, and they may find it easier to tell which molecules is likelier to have DRD2 than the other. In the Bradley Terry model, we first define that $\beta_i \in \mathbb{R}$ be the DRD2 probability of SMILES 1 and β_j be the DRD2 probability of SMILES 2 from the neural networks, and let the outcome of a comparison between SMILES (i, j) be determined by $\beta_i - \beta_j$. The **Bradley-Terry model** treats this outcome as an independent Bernoulli random variable with distribution Bernoulli(p_{ij}), where the log-odds corresponding to the probability p_{ij} that SMILES i is better than SMILES j is [22]

$$\log \frac{p_{ij}}{1 - p_{ij}} = \beta_i - \beta_j.$$

Then, we can solve the probability p_{ij} that SMILES 1 is better than SMILES 2 is

$$p_{ij} = \frac{e^{\beta_i - \beta_j}}{1 + e^{\beta_i - \beta_j}} = \frac{1}{1 + e^{-(\beta_i - \beta_j)}} = Sigmoid(\beta_i - \beta_j)$$

$$p_{ij} = Sigmoid(\text{DRD2 proba SMILES 1} - \text{DRD2 proba SMILES 2})$$

The details of the feedback for Bradley-Terry feedback model is

Inputs: Morgan fingerprints of 2 different SMILES.

Outputs:: Probability p_{ij} that the first SMILES is better than the second SMILES in having DRD2.

Individual feedback: The feedback is 1 if 1st SMILES is better than 2nd SMILES (bradley-terry output > 0.5) else 0, which means 1st SMILES is worse than 2nd SMILES

Batch feedback for REINVENT and acquisition evaluation:

1. Obtaining $P_2^{\text{batch size}}$ permutations of pairs from <batch_size> number of SMILES. This ensures that the relative comparison for all SMILES is comprehensive. For example, SMILES that are better than most other SMILES may imply they have high DRD2 probability.

2. Calculate preference score of SMILES 1 against SMILES 2. For each comparison, individual feedback (0 or 1) for the first SMILES are aggregated.

3. Return the average aggregated score for each SMILES, which is the total

sum above divided by <batch_size> - 1

Loss function to train Pytorch model: BCELoss.

Ranking model, using ListNet architecture

ListNet is a listwise approach for learning to rank, which aims to directly optimize the ranking of a list of items rather than individual pairs like Bradle-Terry model [23]. Its motivation may come from the idea that the comparing model can be quite limited since it is take-all or lose-all feedback, and the ranking model can provide a more neutral ratings between molecules than directly comparing.

In general, the ListNet architecture consists of the following components:

- Input representation: Each SMILES is represented using feature vectors (Morgan fingerprints)
- Softmax function: To convert the raw scores produced by the neural network into probabilities, ListNet uses a softmax function, which ensures that the probabilities of all items in the list sum to one, providing a normalized ranking distribution.
- Loss function: The ListNet uses the Kullback-Leibler Divergence (KLDivLoss) to measure the difference between the predicted ranking distribution and the ground truth ranking distribution.

The details of the feedback for ranking ListNet feedback model is

Inputs: Morgan fingerprints of 3 different SMILES.
Outputs: Softmax preference scores of 3 SMILES w.r.t DRD2 probability.
Individual feedback: SMILES with lowest softmax score receives rank 0, second highest rank 1, and highest one rank 2. Then, they are normalized to [0, 0.5, 1]
Batch feedback for REINVENT and acquisition evaluation:

1. Obtaining $C_3^{\text{batch size}}$ combinations of sets of 3 SMILES from REINVENT output. Due to the exponential increasing nature of combination, the batch size should not be too large. A reasonable value for batch size is 64.
2. Calculate preference scores for 3 SMILES, then convert to ranks [0, 1, 2] and normalize to [0.0, 0.5, 1.0]. Scores are then aggregated.
3. Return the average aggregated score for each SMILES by dividing by $C_2^{\text{batch size}-1}$, which is the number of times each SMILES appear in all combinations.

Loss function: KLDivLoss (Kullback-Leibler divergence), which has formula

$$\mathrm{KL}(P \parallel Q) = \sum_i P(i) \log \left(\frac{P(i)}{Q(i)} \right)$$

3.3 Active learning: acquisition functions

Active learning is a semi-supervised learning as it uses both labeled and unlabeled data. New samples get annotated in an iterative process, where a query strategy is used to choose an example to be queried, and once labeled by an oracle, will result in a model accuracy increment [24].

In this project, there are three query strategies, otherwise known as acquisition functions, to choose most promising SMILES from REINVENT's generated molecules for labelling by the Oracle. They are random, uncertainty and greedy acquisition. All acquisition functions must use batch feedback from the ML model instead of directly using DRD2 probability.

Random acquisition: it selects molecules randomly from the pool of unselected molecules, $U_{r_{best}}$. This approach does not consider the DRD2 probabilities or any feedback but relies on randomness to ensure diverse selection of molecules

Uncertainty acquisition: it selects molecules for which the ML model has the least confidence in its predictions. The usual method for determining this is to choose molecules whose projected scores are closest to 0.5, which denotes maximum uncertainty. Molecules with DRD2 probability close to 0.5 are used for scoring models. In comparing model, pairings of molecules with scores near 0.5 are chosen, which indicate that the model is unsure which is superior. For ranking model, the SMILES with normalized ranking at around the second position (0.5) are chosen, which indicates the ML model's uncertainty of the exact ordering.

Greedy acquisition: it selects molecules that have the highest predicted DRD2 probabilities based on the feedback from the ML model. This approach aims to maximize the DRD2 probability by always choosing the top-performing molecules. In scoring model, this involves selecting the molecules with the highest DRD2 probabilities. In comparing model and ranking model, they aim to maximize the batch feedback score for selected SMILES.

All of the acquisition functions would be tested along with two level of Oracle's labelling noise: 0.0 and 0.1. The noise of 0.0 means near perfect labelling of DRD2 proability, while noise of 0.1 means the labelling is not entirely accurate

but generally following in correct direction. With 3 feedback models, 3 acquisition functions and 2 noise levels, there are in total $3 \times 3 \times 2 = 18$ different running cases. They would all be benchmarked in the result section.

3.4 Novelty, diversity, SA, and QED score

As per objectives of de novo molecular design, the novelty score, diversity score, synthetic accessibility (SA) score, and quantitative estimate of drug-likeness (QED) score are important metrics to assess the quality of generated molecules. All metrics areessential to guarantee that the molecules produced are not only distinct but also practical to synthesis and likely to have drug-like qualities.

- **Novelty score**
  - **Definition**: The novelty score measures the fraction of the generated molecules not present in the training set [25]. It has range [0, 1].
  - **Range meaning**: Low novelty scores indicates overfitting and high novelty means the generating model can discovers new structures [25]
- **Diversity score**
  - **Definition**: The diversity score measures the internal chemical diversity within the generated molecules set G [26]. It has range [0, 1].

    $$\text{IntDiv}_\text{p}(G) = 1 - \sqrt[p]{\frac{1}{|G|^2} \sum_{m_1, m_2 \in G} Tanimoto(m_1, m_2)^p}.$$

    This metric helps detect mode collapse, which means the model produces samples in a limited chemical space and ignore other spaces [25].
  - **Range meaning**: A higher value of internal diversity indicates larger area of chemical space covered in the generated set and vice versa
- **Synthetic accessibility (SA) score**
  - **Definition**: The SA score evaluates how easily a molecule can be synthesized, ranging from 1 (very easy) to 10 (very difficult).
  - **Range meaning**: Molecules with SA scores between 1 and 3 are considered preferable as they are more likely to be synthesized efficiently in a laboratory setting [27].
- **QED score (Quantitative estimate of drug-likeness)**

- **Definition**: The QED score is a composite metric that evaluates the drug-likeness of a compound based on several molecular properties, ranging from 0 to 1.
- **Range meaning**: A QED score above 0.5 is considered the threshold for drug-likeness, making it a useful filter in drug discovery [28].

3.5 Molecular descriptor filters

Previously, it is mentioned that besides DRD2 maximization, we also aim to generate drug-like molecules based on molecular descriptors, which heavily influence pharmacokinetic parameters. In general, pharmacokinetic (PK) parameters play a crucial role in understanding how a drug interacts with the body during drug development. These parameters provide details into various aspects of a drug when it travels through the body, known as ADME: Absorption, Distribution, Metabolism, and Excretion [29][30]. They include Cmax (highest plasma concentration of a drug after administration), Tmax (time to maximum concentration), half-life (time it takes for the drug concentration in the body to decrease by half), AUC (total exposure to a drug over time), volume of distribution (how extensively a drug distributes into tissues relative to its concentration in plasma), and clearance (rate at which the drug is removed from the body) [31]. Usually, the properties that affect these PK parameters are the molecular descriptors, which are used to characterize the physical, chemical, and structural properties of molecules. Together, they can be called filters, since only drugs that bypass these molecular descriptor filters' threshold should be screened for further evaluation. This cuts down screening time and improves drug-likeness of molecules.

- **LogP (Partition coefficient)**
  - **Definition**: LogP quantifies the lipophilicity of chemical compounds, which influences their absorption, distribution, and overall pharmacokinetic properties.
  - **Filter range**: Compounds with a logP between 1 and 5 are typically preferred, as they are more likely to have favorable pharmacokinetic profiles [32].
- **Molecular weight**
  - **Definition**: This is the mass of a molecule measured in Daltons.

  - **Filter range**: Molecules with a molecular weight under 500 Daltons are generally considered ideal for oral drug candidates as per Lipinski's rule of five [32].

- **Hydrogen bond donors (H-donors) and acceptors (H-acceptors)**
  - **Definition**: H-donors are atoms in a molecule that can donate hydrogen bonds, whereas H-acceptors are atoms that can accept hydrogen bonds.
  - **Filter range**: According to Lipinski's rule of five, having no more than 5 hydrogen bond donors and no more than 10 hydrogen bond acceptors is preferred for good bioavailability [32].

- **TPSA (Topological Polar Surface Area)**
  - **Definition**: TPSA is the surface area of a molecule that is polar.
  - **Filter range**: Molecules with a TPSA of 140 Å² or less are more likely to have good oral bioavailability, as per Veber's rule [33].

- **Number of Rotatable Bonds**
  - **Definition**: This is the number of bonds in a molecule that can rotate freely.
  - **Filter range**: According to Veber's rule, having 10 or fewer rotatable bonds is ideal for maintaining good oral bioavailability [33].

- **Number of Rings**
  - **Definition**: This is the number of ring structures within a molecule.
  - **Filter range**: The presence of ring structures can impact the rigidity and overall stability of a molecule. Muegge's rule suggests that having up to 7 rings is favorable for drug-like properties [34].

The table below summarizes all objectives and filters used in this project

	Description	Lower	Higher	Known rules
DRD2 probability	Objective to maximize	0.75	1.0	0.75 is average DRD2 probability of SMILES actually having DRD2 predicted by TDC Oracle
Molecule weight	Filtering	0	500.0	Lipinski's rule of five [32]
Hydrogen bond donors number	Filtering	0	5	Lipinski's rule of five [32]
Hydrogen bond acceptors number	Filtering	0	10	Lipinski's rule of five [32]
TPSA	Filtering	0.0	140.0	Veber et al. [33]
Number of rotatable bonds	Filtering	0	10	Veber et al. [33]
Number of rings	Filtering	0	7	Muegge et al. [34]

4 Results

4.1 Performance metrics evolution between running cases

The ROC (Receiver Operating Characteristic) curve and AUC (Area Under the Curve) are used to evaluate the performance of binary classification for DRD2 by plotting the true positive rate against the false positive rate at various thresholds of DRD2 probability. The AUC gives a single scalar value to quantify the overall performance of the classifier. Random classifier has AUC = 0.5 and the better the classifier, the higher the AUC score.

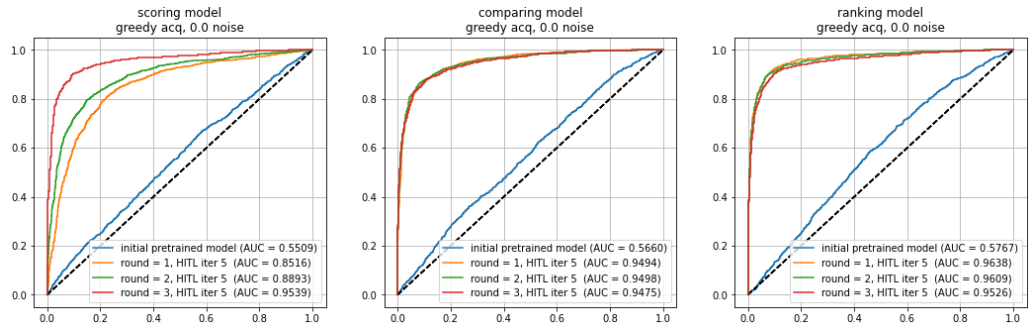

Figure 2. *ROC curve improvement after 3 REINVENT rounds*

Initially, the pre-trained models show relatively low AUC values, which imitate a human component, as they have limited ability to tell apart molecules with and without DRD2. However, with each round of 5 HITL iterations, there is a considerable improvement in AUC values for all models.

Across all running cases, the AUC values indicate that all models perform well as expected, with minimal differences between them. The ranking model consistently achieves the highest AUC values, followed closely by the comparing model and the scoring model. This suggests that while all three models are effective in predicting

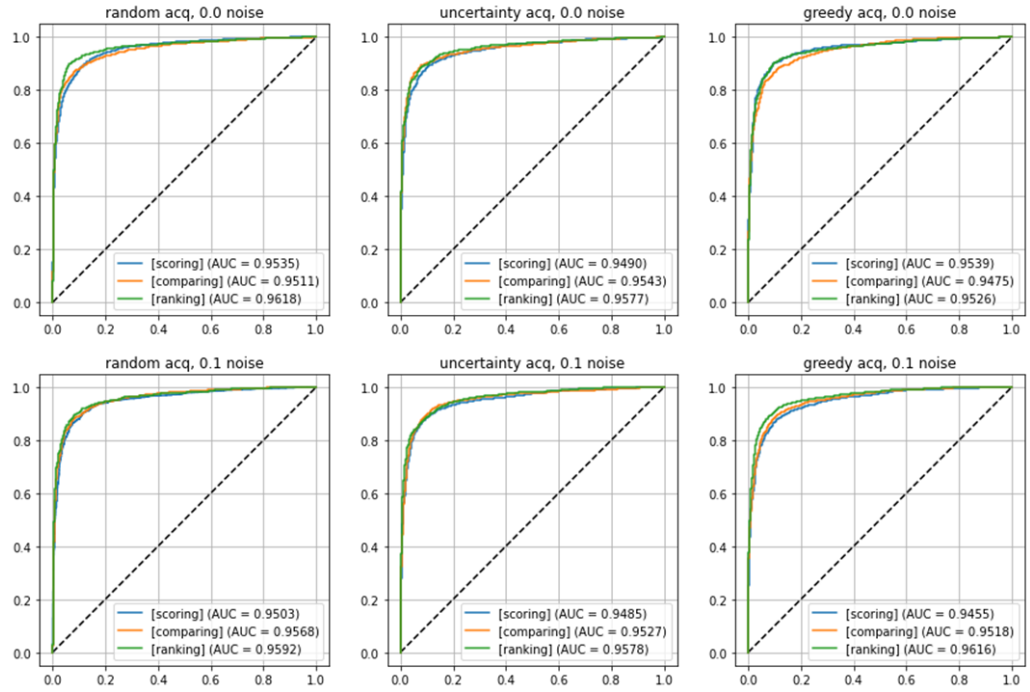

Figure 3. *ROC curve for DRD2 classification at last REINVENT round, HITL iteration 5*

DRD2 activity, the ranking model slightly outperforms the others, providing the most reliable prediction of DRD2 presence.

4.2 Performance metrics comparison between running cases

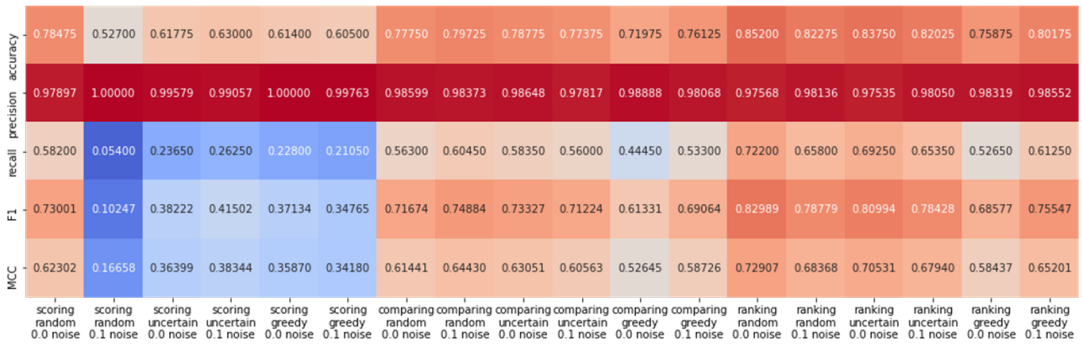

Figure 4. *Performance metrics of the last ML model for 18 running cases*

These metrics are obtained by evaluating the models on a testing dataset consisting of 1000 smiles with DRD2 and also 1000 without DRD2. As a result, this dataset is balanced and large enough to reliably tell the model's performance. Looking at the heatmap, we can observe that most models have high precision, which is meaningless because the positive case (DRD2) is much rarer during the training process. As a result, we should focus on finding the configuration that has highest recall. It is apparent that the ranking model consistently has higher recall than comparing model, which in turn has higher recall than scoring

model. It is surprising since comparing and ranking molecules can deliver better classification result than directly learning the DRD2 probability. Specifically, the best configuration is the ranking model that uses random acquisition without human noise, with highest accuracy, recall, F1 and MCC score. This suggests that diversity of SMILES in training the model improves DRD2 presence classification.

4.3 Percentage of filtered generated molecules

Figure 5. *Percentage of molecules combined from all rounds that pass through each filter*

This table shows the percentage of molecules that satisfy the lower and higher thresholds molecule descriptor. We can observe that most molecules generated by REINVENT already satisfy the filters except logP and number of hydrogen bond donors. This can probably indicate that the constraints on logP and number of hydrogen bonds constraint are less common for drug-like molecules. Regarding drd2 probability, it is apparent that the scoring model is the most capable of generating molecules with DRD2, followed by the comparing model and lastly by the ranking model. This is the major advantage of the baseline scoring model when we need to directly optimize some properties instead of focusing on novelty or diversity. As a result, after passing through all filters, the scoring model tends to have larger percentage of filtered smiles, followed by the comparing model and lastly the ranking model.

4.4 DRD2 probability of generated molecules

Previously, we have observed that the scoring model has the highest percentage of SMILES passing through all filters, but it does not imply that the scoring model in general generates molecules with highest DRD2 probability. That is why it is necessary to compare the DRD2 probability before and after filtering, which are labeled "unfiltered" and "filtered" in the figure below.

Coincidentally, the scoring model also happens to generate molecules with highest DRD2 probability, as the mean value and 75 percentile are consistently higher than those of comparing and ranking models by a large margin in the figure below. The comparing model tends to generate molecules with higher DRD2 probability than the comparing model, which is more apparent in the unfiltered version but less clear in the filtered version. Overall, we can conclude that in terms of generating molecules with DRD2 presence, the scoring model is the best, followed by the comparing model and lastly the ranking model. This is true for both unfiltered and filtered SMILES cases.

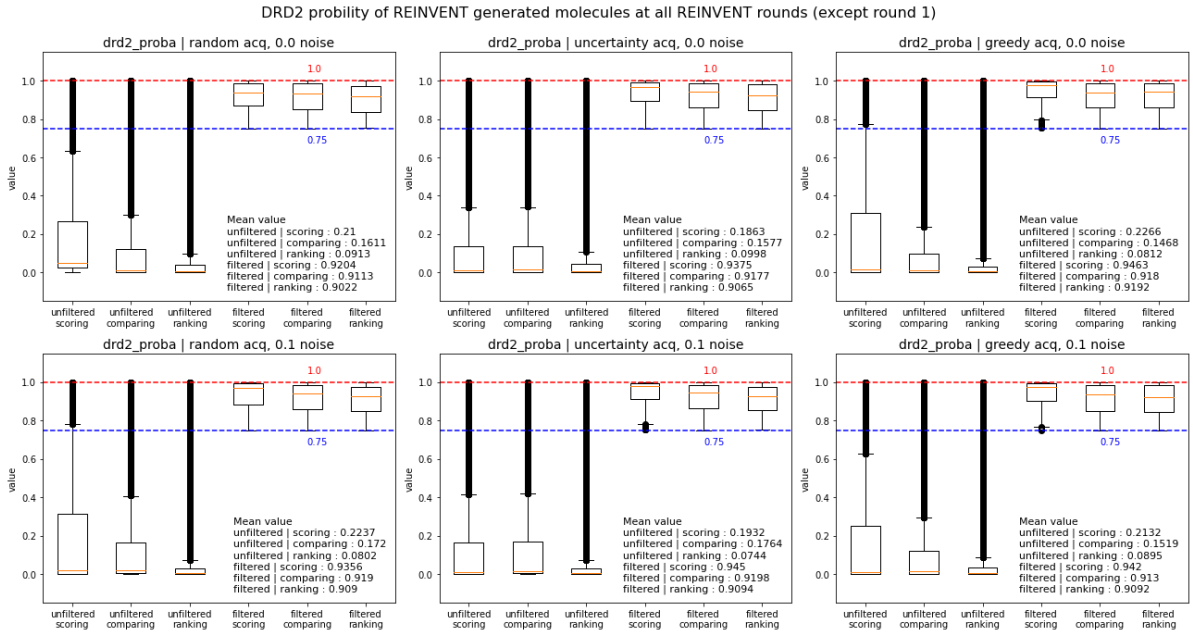

Figure 6. *DRD2 probability of generated molecules combined from all rounds*

4.5 Benchmark scores of generated molecules

The result before should make it clear about which model should be preferred when targeting chemical properties. It is time to consider novelty, diversity, synthesizability and drug-likeness to provide a fairer picture for 3 feedback models.

Regarding novelty score, it is hard to determine which model is highly capable

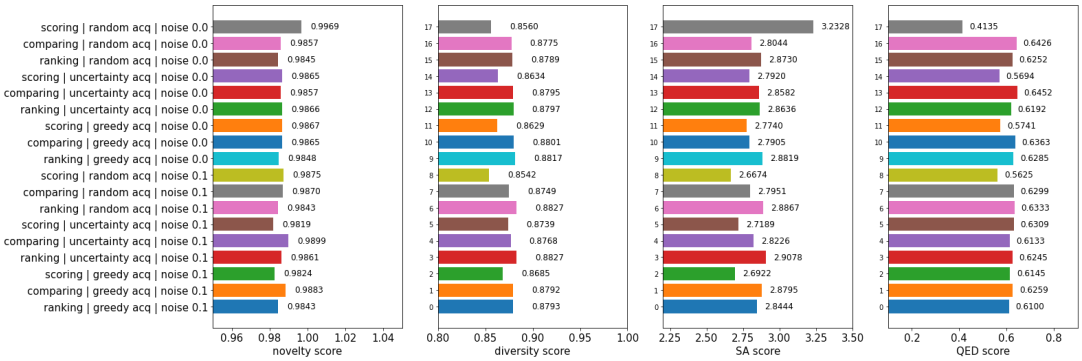

Figure 7. *Benchmark scores of unfiltered generated molecules combined from all rounds*

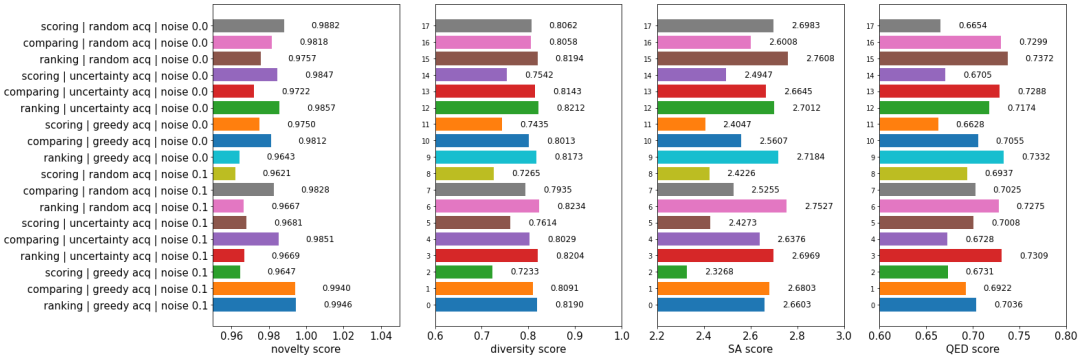

Figure 8. *Benchmark scores of filtered generated molecules combined from all rounds*

of generating the most novel molecules, but it appears that the comparing model has the highest novelty score for 5 out of 12 configurations in both filtered and unfiltered versions, followed by 4 out of 12 for scoring model and 3 out of 12 for ranking model. However, we should note that this comparison is not guaranteed to be true if the three models are tested with different running settings.

Regarding diversity score, we have a major discovery here, where the ranking model consistently outperforms others in all 12 configurations in both filtered and unfiltered versions. This strongly suggest that ranking feedback mechanism is much more versatile to encourage generative models like REINVENT to explore larger chemical spaces that are likely to possess the targeted chemical properties. We also conclude that the comparing model, while always has lower diversity score than comparing model, consistently outperforms the scoring model.

Regarding SA score, we first note that molecules with low SA scores generally have several common characteristics. They often contain simple building blocks and have have fewer rings, stereocenters, and chiral centers, as well as a lower overall molecular weight. They also tend to lack highly strained or unusual chemical structures and have a small number of functional groups that are hard to manipulate [27]. We have an additional discovery from the two graphs, where

the scoring model consistently has lowest SA score for 10 out of 12 configuratons. This suggest that due to the lack of diversity as mentioned earlier, it is likely to explore a limited chemical space that are likelier to have DRD2. This by-product process makes the molecules generated by scoring feedback more likely to be synthesizable. Then, we see that the comparing model has higher SA score and finally is ranking model, which has highest SA score. This means that SA score is in a way inversely proportional to the diversity score: the more diverse a set of molecules, the harder it is to synthesize all of them.

Regarding QED score, it appears that the comparing model and the ranking model have close competition, where the comparing model tops 5 out of 12 configurations and ranking model tops 6 out of 12 configurations. Interestingly, the comparing model outperforms others mostly in the unfiltered version, while the ranking model outperforms others mostly in the filtered version. This suggests that overall, even though the comparing model can encourage the generative models to discover drug-like molecules, they will be not likely to contain the targeted chemical property or satisfy all molecule descriptor filters. On the other hand, the ranking model not only generates molecules with good targeted chemical property (DRD2), but it also balances them with high QED score as well. To conclude, the ranking model performs best, followed by the comparing model and lastly the scoring model for drug-likeness criterion.

4.6 Examples of best generated molecules

To have a better visualizations of the generated molecules, we can use RDKit to plot the best SMILES from each 18 running case from the filtered SMILES. The best SMILES is determined by first choosing the top 20 SMILES with highest QED score, then we sort them by SA score and choose the top 10 SMILES, and finally we sort the 10 SMILES by their DRD2 probability and fetch the top one.

Generally speaking, the molecules produced by the scoring model have simple structures with few branches and fewer functional groups. Because it produces molecules that are often simpler (as demonstrated before) and more focused on specific chemical properties known to be related with DRD2 activity, it directly predicts the likelihood that a given SMILES string contains DRD2 activity. A design bias towards high predictability and ease of synthesis can be seen by the compounds' common pharmacophores and synthetic routes.

In contrast to the scoring model, the molecules produced by the comparing model

show greater structural complexity and diversity. The different ring systems and functional groups demonstrate this diversity and show how well the model is able to distinguish between the relative benefits of various chemical configurations.

Finally, the ranking model produces molecules that maintain a balance between drug-likeness, diversity, and novelty. These molecules are more oriented toward reaching a high rank across various metrics, and they frequently exhibit a combination of properties from the other SMILES generated by scoring and comparing models. The structures produced by ranking feedback display a combination of diverse functional groups and complex ring systems, indicating that the model successfully incorporates various feedbacks to rank molecules according to DRD2 probability as well as their drug-likeness.

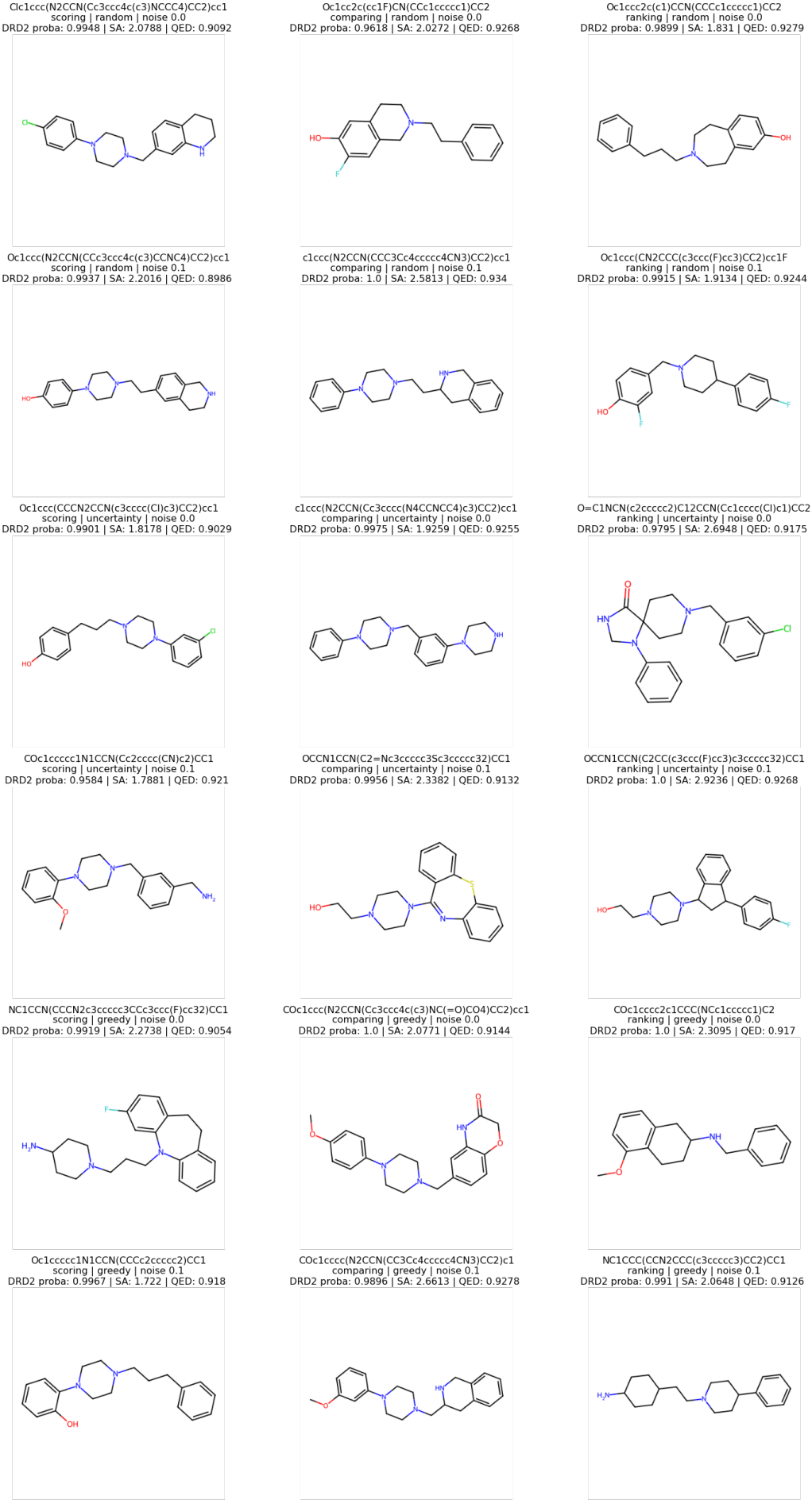

Figure 9. *Molecules generated by different feedback models*

5 Discussion

5.1 Feedback model comparison

In brief, the ranking model (ListNet) consistently excels in DRD2 classification with high MCC score, passing molecular descriptors, maximizing diversity, and achieving high QED scores. On the other hand, the scoring model shows the best performance in generating molecules that maximize SA score, DRD2 probability and pass the DRD2 probability threshold. The comparing model (Bradley-Terry) keeps a balance between these two models, performing moderately well across all criteria, particularly excelling in generating molecules with high novelty score. We can summarize the performance of the three feedback models in the table below, where rank 1 means best and rank 3 means worst.

	scoring model	**comparing model Bradley Terry**	**ranking model ListNet**
ML model with highest MCC on DRD2 classification	3	2	1
Generate molecules that are likely to pass molecular descriptors	3	2	1
Generate molecules that are likely to pass DRD2 probability threshold	1	2	3
Generate molecules that maximize DRD2 probability	1	2	3
Generate molecules that maximize novelty score	2	1	3
Generate molecules that maximize diversity score	3	2	1
Generate molecules that minimize SA score	1	2	3
Generate molecules that maximize QED score	3	2	1

5.2 Acquisition function comparison

Like the previous approach, it is possible to also measure the rankings of the acquisition functions as well. The findings and rankings are only a general trend, and naturally there are result variations within each ranking.

The random acquisition function has the highest MCC for DRD2 classification but the lowest for producing molecules that would pass molecular descriptors and DRD2 probability thresholds. It excels at maximizing novelty and diversity scores,

showing higher ability to explore more chemical spaces.

Uncertainty acquisition balances performance by placing second in most categories, generating varied and new molecules while retaining intermediate scores for molecular descriptor and DRD2 probability criteria.

Greedy acquisition is the best at generating molecules that pass molecular descriptors and DRD2 probability thresholds, as well as maximizing DRD2 probability and minimizing SA scores, but it ranks last in terms of novelty and diversity, focusing more on known parameter optimization.

	random acquistion	**uncertainty acquisition**	**greedy acquisition**
ML model with highest MCC on DRD2 classification	1	2	3
Generate molecules that are likely to pass molecular descriptors	3	2	1
Generate molecules that are likely to pass DRD2 probability threshold	3	2	1
Generate molecules that maximize DRD2 probability	3	2	1
Generate molecules that maximize novelty score	1	2	3
Generate molecules that maximize diversity score	1	2	3
Generate molecules that minimize SA score	3	2	1
Generate molecules that maximize QED score	1	2	3

5.3 Human noise comparison

Finally, we can assess the performance of models at noise levels of 0.0 and 0.1 across previous metrics. Models with a noise level of 0.1 regularly outperforms the other with a noise level of 0.0 in terms of producing molecules that pass molecular descriptors while maximizing novelty, diversity, and QED scores. However, models with noise level 0.0 outperform othe other one in terms of producing compounds that maximize DRD2 probability while minimizing SA scores. Which noise is better for passing DRD2 probability thresholds and getting the highest MCC in DRD2 classification remain questionable. This shows that introducing noise can improve some elements of molecule production (targeting novelty and diversity) while reducing criteria like synthesizability.

	noise 0.0	noise 0.1
ML model with highest MCC on DRD2 classification	uncertain	uncertain
Generate molecules that are likely to pass molecular descriptors	2	1
Generate molecules that are likely to pass DRD2 probability threshold	uncertain	uncertain
Generate molecules that maximize DRD2 probability	1	2
Generate molecules that maximize novelty score	2	1
Generate molecules that maximize diversity score	2	1
Generate molecules that minimize SA score	1	2
Generate molecules that maximize QED score	2	1

6 Conclusion

When working with de novo molecular design, the feedback mechanism, acquisition function and the noise in labelling offer a wide variety of customization. In other words, depending on the objective of the generation task, we can use a specific configuration to ensure that it will be likely to achieve quantifiable results.

To focus on generating molecules that directly targets chemical properties as much as possible like DRD2 and easiest synthesizability, the scoring model used with the greedy acquisition function without labelling noise is a good option to experiment initially. The scoring model is preferable when computational resource is limited.

To focus on generating novel molecules while keeping a balance across all metrics, the comparing model used with the uncertainty acquisition function is a good option to start with. The comparing model however needs a little more computational resource than the scoring model.

To focus on generating diverse and drug-like molecules as much as possible, the ranking model used with the random acquisition and noisy labelling is possibly the best combination to experiment with. However, the ranking model requires extensive computational resource during an optimization step of REINVENT since all sets of 3 molecules (combinatorial) are computed to infer their preferences.

In conclusion, this research highlights the impact of several feedback models on de novo molecular design with REINVENT. We used scoring, comparing, and ranking

models to generate compounds with high DRD2 likelihood, novelty, diversity, synthesizability and drug-likeness criteria, with variable degrees of success. Future research could further refine these feedback models or introduce new feedback mechanisms, such as binary preference (yes/no liking given a SMILES), multiple rankings (4-5 SMILES ranking), or even multiple binary preference (yes/no liking given k-criteria given a SMILES).

7 Appendix

The project code for this research project can be found from this Github repository: https://github.com/SpringNuance/Human-In-The-Loop-De-Novo-Molecular-Design

In order to run the project code at Base-Code-Binh, please install the environment of ReinventCommunity located at Reinvent-Community-original/environment.yml. For the author, using any other yaml environment is likely to result in compatibility issues. Additionally, this project is conducted for REINVENT version 3.2. Latest REINVENT version 4.0 has been released, which may not be compatible with the author's source code.

References

[1] S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang, and S. H. Bryant, "Pubchem substance and compound databases," *Nucleic Acids Research*, vol. 44, no. D1, pp. D1202–D1213, 2016.

[2] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, and G. R. Hutchison, "Open babel: An open chemical toolbox," *Journal of Cheminformatics*, vol. 3, p. 37, 2011.

[3] D. Weininger, "Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules," *Journal of Chemical Information and Computer Sciences*, vol. 28, no. 1, pp. 31–36, 1988.

[4] C. A. James, D. Weininger, and J. Delany, *Daylight Theory Manual*. Daylight Chemical Information Systems, Inc., 2007. Link: `https://www.daylight.com/dayhtml/doc/theory/theory.smarts.html`.

[5] V. Gillet, *De Novo Molecular Design*, vol. 4. 2000.

[6] H.-H. Hsu, C.-H. Huang, and S.-T. Lin, "New data structure for computational molecular design with atomic or fragment resolution," *Journal of Chemical Information and Modeling*, vol. 59, no. 8, pp. 3325–3335, 2019.

[7] A. Tharwat and W. Schenck, "A survey on active learning: State-of-the-art, practical challenges and research directions," *Mathematics*, vol. 11, no. 4, 2023.

[8] T. Kaufmann, P. Weng, V. Bengs, and E. Hüllermeier, "A survey of reinforcement learning from human feedback," 2023.

[9] S. Takeda, H. Kaneko, and K. Funatsu, "Chemical-space-based de novo design method to generate drug-like molecules," *Journal of Chemical Information and Modeling*, vol. 56, no. 9, pp. 1667–1681, 2016.

[10] S. K. Dhanda, D. Singla, A. Mondal, and G. Raghava, "Drugmint: a webserver for predicting and designing of drug-like molecules," *Biology Direct*, vol. 8, p. 28, 2013.

[11] Y. Liang, R. Zhang, L. Zhang, and P. Xie, "Drugchat: Towards enabling chatgpt-like capabilities on drug molecule graphs," *arXiv preprint arXiv:2309.03907*, 2023.

[12] T. Blaschke, J. Arús-Pous, H. Chen, C. Margreitter, C. Tyrchan, O. Engkvist, K. Papadopoulos, and A. Patronov, "Reinvent 2.0: An ai tool for de novo drug design," *Journal of Chemical Information and Modeling*, vol. 60, no. 9, pp. 4423–4433, 2020.

[13] H. H. Loeffler, J. He, A. Tibo, J. Janet, A. Voronov, L. H. Mervin, and O. Engkvist, "Reinvent 4: Modern ai-driven generative molecule design," *Journal of Cheminformatics*, vol. 16, no. 1, p. 812, 2024.

[14] T. Blaschke, J. Arús-Pous, H. Chen, C. Margreitter, C. Tyrchan, O. Engkvist, K. Papadopoulos, and A. Patronov, "Supporting information reinvent 2.0: An ai tool for de novo drug design," *ChemRxiv*, 2020.

[15] W. Zhang, M. Lei, Q. Wen, D. Zhang, G. Qin, J. Zhou, and L. Chen, "Dopamine receptor d2 regulates glua1-containing ampa receptor trafficking and central sensitization through the pi3k signaling pathway in a male rat model of chronic migraine," *Journal of Headache and Pain*, vol. 23, no. 1, p. 45, 2022.

[16] S. Pierce, Z. Fang, Y. Yin, L. West, M. Asher, T. Hao, X. Zhang, K. Tucker, A. Staley, Y. Fan, W. Sun, D. Moore, C. Xu, Y.-H. Tsai, J. Parker, V. Prabhu, J. Allen, D. P. Lee, C. Zhou, and V. Bae-Jump, "Targeting dopamine receptor d2 as a novel therapeutic strategy in endometrial cancer," *Journal of Experimental Clinical Cancer Research*, vol. 39, no. 1, p. 38, 2020.

[17] H. Lee, S. Shim, J. Kong, M.-J. Kim, S. Park, S.-S. Lee, and A. Kim, "Overexpression of dopamine receptor d2 promotes colorectal cancer progression by activating the ŏ3b2-catenin/zeb1 axis," *Cancer Science*, vol. 112, no. 4, pp. 1503–1515, 2021.

[18] M. Wang, Z. Wu, J. Wang, G. Weng, Y. Kang, P. Pan, D. Li, Y. Deng, X. Yao, Z. Bing, C.-Y. Hsieh, and T. Hou, "Genetic algorithm-based receptor ligand: A genetic algorithm-guided generative model to boost the novelty and drug-likeness of molecules in a sampling chemical space," *Journal of Chemical Information and Modeling*, vol. 64, no. 2, pp. 326–345, 2024.

[19] J. Fotie, C. M. Matherne, and J. E. Wroblewski, "Silicon switch: Carbon-silicon bioisosteric replacement as a strategy to modulate the selectivity, physicochemical, and drug-like properties in anticancer pharmacophores," *Chemical Biology and Drug Design*, vol. 91, no. 4, pp. 239–249, 2023.

[20] I. Sundin, A. Voronov, H. Xiao, K. Papadopoulos, E. J. Bjerrum, M. Heinonen, A. Patronov, S. Kaski, and O. Engkvist, "Human-in-the-loop assisted de novo molecular design," *ChemRxiv*, 2022.

[21] I. Sundin, A. Voronov, H. Xiao, K. Papadopoulos, E. Bjerrum, M. Heinonen, A. Patronov, S. Kaski, and O. Engkvist, "Human-in-the-loop assisted de novo molecular design," *Journal of Cheminformatics*, vol. 14, 12 2022.

[22] Unknown, "Lecture 24: Bradley-terry model," 2017. Accessed: 2024-07-16.

[23] Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, and H. Li, "Learning to rank: from pairwise approach to listwise approach," in *Proceedings of the 24th international conference on Machine learning*, pp. 129–136, ACM, 2007.

[24] E. Mosqueira-Rey, E. Hernández-Pereira, D. Alonso-Ríos, J. Bobes-Bascarán, and Fernández-Leal, "Human-in-the-loop machine learning: a state of the art," *Artificial Intelligence Review*, vol. 56, 08 2022.

[25] D. Polykovskiy, A. Zhebrak, B. Sanchez-Lengeling, S. Golovanov, O. Tatanov, S. Belyaev, R. Kurbanov, A. Artamonov, V. Aladinskiy, M. Veselov, A. Kadurin, S. Johansson, H. Chen, S. Nikolenko, A. Aspuru-Guzik, and A. Zhavoronkov, "Molecular sets (moses): A benchmarking platform for molecular generation models," *arXiv preprint arXiv:1811.12823*, 2018.

[26] M. Benhenda, "Chemgan challenge for drug discovery: can ai reproduce natural chemical diversity?," *arXiv preprint arXiv:1708.08227*, 2017.

[27] P. Ertl and A. Schuffenhauer, “Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions,” *Journal of Cheminformatics*, vol. 1, p. 8, 2009.

[28] G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan, and A. L. Hopkins, “Quantifying the chemical beauty of drugs,” *Nature Chemistry*, vol. 4, pp. 90–98, 2012.

[29] M. Rasool and S. Läär, “Development and evaluation of a physiologically based pharmacokinetic model to predict carvedilol-paroxetine metabolic drug-drug interaction in healthy adults and its extrapolation to virtual chronic heart failure patients for dose optimization,” *Expert Opinion on Drug Metabolism Toxicology*, vol. 17, no. 6, pp. 695–709, 2021.

[30] K. Lin, J. Tibbitts, and B.-Q. Shen, *Pharmacokinetics and ADME characterizations of antibody-drug conjugates*, vol. 1045 of *Methods in Molecular Biology*, pp. 117–131. Humana Press, 2013.

[31] A. I. Omar and K. Na-Bangchang, “Pharmacokinetic studies of nanoparticles as a delivery system for conventional drugs and herb-derived compounds for cancer therapy: a systematic review,” *International Journal of Nanomedicine*, vol. 14, pp. 9159–9173, 2019.

[32] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, “Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings,” *Advanced Drug Delivery Reviews*, vol. 23, no. 1-3, pp. 3–25, 1997.

[33] D. F. Veber, S. R. Johnson, H. Y. Cheng, B. R. Smith, K. W. Ward, and K. D. Kopple, “Molecular properties that influence the oral bioavailability of drug candidates,” *Journal of Medicinal Chemistry*, vol. 45, no. 12, pp. 2615–2623, 2002.

[34] I. Muegge, S. L. Heald, and D. Brittelli, “Simple selection criteria for drug-like chemical matter,” *Journal of Medicinal Chemistry*, vol. 44, no. 12, pp. 1841–1846, 2001.