Question to Yasmine:

1. **What key points should I emphasize in my presentation to meet the supervisor's expectations?**

   He's more into methodology descriptions (so you should describe the three feedback models) and « why you're doing what you're doing ».
   And as I suggested in my feedback: you should emphasize that the project has the motivation to push the boundaries of the state of the art in HITL assisted molecular design by proposing new feedback mechanisms to give feedback on generated molecules and provide recommendations for users of generative chemistry of which feedback mechanism works best to improve which criteria (that's already what you're telling in your conclusions).
   Most important thing is the Motivation, why we are doing this project, and describes very carefully the 3 feedback models
   Also, your project is an applied project not a methodology one, so emphasize that at the beginning of the presentation so that Sami gets the context.

   Motivation: improve the human in the loop assisted de novo molecular design. This framework already exists, but my contribution is to invent different modes of human interaction like pairwise and ranking than the binary like, dislike feedback, to see which user feedback is more efficient in learning chemist preference about the molecules.
   To target physical properties in drug design, or many other aspects like having preferable molecular descriptors or having favorable quantitative scores like SA score.

2. **How long should the speaking part be, and how long should the QA be?:**
   Presentation: 40 minutes, QA: 20 minutes

3. **Do we have a link to the final presentation meeting already?:**
   https://aalto.zoom.us/j/64549731963

4. **What type of questions can I expect from the supervisor, and how should I prepare for them?**

   Focus very carefully on the Motivation (why are you doing this project) and the methodologies. For other parts, your results and discussions can be not perfect, but for methodologies and motivation, you have to defend them very well, and professor is very likely to ask about motivation and methodologies

10 seconds = 20 words
30 seconds = 60 words
60 seconds = 120 words
75 seconds = 150 words
90 seconds = 180 words
120 seconds = 240 words
150 seconds = 270 words

Presentation scripts:

Slide 1: Introduction (30 seconds)

Good morning, Yasmine, and professor Samuel. So today, I am going to present my final research project on human-in-the-loop reinforcement learning for exploring innovative feedback mechanisms in de novo molecular design. For disclaimer, this project is an applied research project instead of a methodological focused project. I hope that this will make clear the context of my work from this point.

Slide 2: Table of contents (30 seconds)

Let's quickly walk through the contents. We'll start with some literature review on HITL- Machine Learning for de novo drug design. Then, I will discuss the project motivation and objectives, followed by a review of REINVENT software and three user feedback mechanisms. I will explore the HITL workflow, present the results, and discuss these findings. We'll conclude with some general advices on user feedback models.

Slide 3: State of the art HITL ML for de novo drug design (10 seconds)

Let's explore how current advances in HITL machine learning are improving the quality and reliability of de novo drug design

Slide 4: De-Novo Molecular Design (75 seconds)

Alright, let's dive into what 'de-novo' molecular design really means. The term 'de novo' is in Latin for 'from the beginning'. This concept is all about starting from scratch. We're not just improving on existing molecules. Instead, we're creating entirely new molecular structures using computational methods.

So, how do we do this? Well, we use algorithms to search through vast chemical spaces, exploring and optimizing molecules based on predefined criteria such as biological activity, how drug-like they are, and their synthesizability. There is no guarantee of finding the best ones as the chemical space is astronomically large, but sometimes we can end up with potentially groundbreaking therapeutic drugs.

And there's a variety of approaches here—from using atom-based methods, where we build the molecule atom by atom, to fragment-based methods, where we piece together larger parts of molecules like a puzzle. This flexibility lets researchers find molecules that can target specific therapeutic qualities.

Slide 5: Human-in-the-loop De-Novo molecular design (120 seconds)

In the realm of de-novo molecular design, the HITL framework plays a crucial role by integrating human domain expertise directly into the modeling process. This connection between computer and cognitive science allows for more refined and targeted drug design.

The concept of reinforcement learning from human feedback, or RLHF, introduces a vital human component into the loop, which allows experts to directly define the objectives of the modeling process. This integration ensures that the development of new drugs aligns more closely with practical needs and chemist's preference.

Several notable projects show the advancements in this field. For example, Robin 2020 developed grünifai, an interactive platform that optimizes small molecules for drug discovery. This system integrates adjustable models, represents chemical spaces continuously, and uses scalable optimization algorithms, all refined by user feedback to balance many molecular properties.

Another significant contribution is from Sundin 2022, who presented a human-in-the-loop machine learning approach using a probabilistic model. This system incorporates active learning to integrate feedback effectively into a multi-parameter optimization scoring function. This would be the backbone referenced paper for this research project.

Furthermore, Choung 2023 applied AI learning-to-rank techniques, using feedback from human chemists to improve the lead optimization process in drug discovery. These examples demonstrate how human insights can significantly aid the computational models.

Slide 6: Where is this research project's domain on this chart? (60 seconds)

This is a diagram of the entire landscape of HITL - ML from Mosqueira 2022 paper, and it is important to know where this research project's domain may fall into. Our project primarily resides within the 'Human-in-the-loop Machine Learning area, as highlighted in red. This framework emphasizes the dynamic integration of human feedback directly into the ML workflow, predicting the model's ability to make decisions and predictions.

Additionally, our project intersects with 'Active Learning' where the ML model actively queries new molecules for labeling by humans. This interaction allows the model to learn more efficiently and become more precise in predicting molecular characteristics such as DRD2 affinity.

Slide 7: Research project motivation and objectives (10 seconds)

Next, we will explore the driving motivation and objective goals behind this research project

Slide 8: Research project motivation (120 seconds) (240 words)

In this project, the primary motivation is to assist the de novo molecular design by integrating diverse HITL feedback mechanisms. We would aim to capture and use chemist intuition more effectively, pushing the boundaries of traditional approaches that does not rely on HITL.

Why is this important? Previous works like Sundin 2022 rely on binary feedback mechanisms—simple likes or dislikes—which can be limiting. Our objective is to explore and develop a broader range of feedback models, specifically the direct scoring, pairwise comparison and ranking user feedbacks. These models are designed to provide different feedback, which may help generative models to capture chemist preferences.

Take the case of targeting the dopamine receptor D2 (DRD2), which plays a crucial role in managing disorders like schizophrenia and Parkinson's disease.

The expected outcome of this project, therefore, is a more efficient molecular design process that uses human expertise to discover molecules that not only have good chemical properties but are also novel, diverse and drug-like. This approach aligns closely with the needs of practitioners in the bioinformatic field, making this research directly applicable to drug discovery

In this project, DRD2 is used as a toy use case in this project to demonstrate the effectiveness of our user feedback models. This allows us to evaluate our results using the existing Oracle model on DRD2 affinity classification.

Slide 9: Research project objectives (90 seconds) (180 words)

In this research, our objectives are structured into three steps

The first step is De Novo Molecular Design, which generates novel molecules that specifically target the DRD2 receptor, using the REINVENT platform. This step is not just about DRD2 affinity but also ensuring that these molecules are diverse and maintain drug-like properties.

The second step is Human-in-the-loop Workflow, which involves developing models that can learn from human feedback effectively. We are using three distinct user feedback — scoring, pairwise comparison, and ranking—to see which provides the best input for our AI systems. These mechanisms are complemented by three different acquisition strategies

The third step is to build a Machine Learning Model to predict DRD2 Affinity. This objective focuses on developing a surrogate machine learning model that predicts the affinity of newly

designed molecules for the DRD2 receptor. This model serves as a benchmark, allowing us to compare the learning performance of our user feedback models against existing standard models, or the 'Oracles', in the field.

Through these objectives, I can show that human chemists are important to the discovery process, making the AI-generated molecules more relevant and have good properties

Slide 10: REINVENT: generative RL model (10 seconds)

Coming up next, we have REINVENT, which is a generative reinforcement learning model used to design new molecules.

Slide 11: REINVENT a brief look (75 seconds)

REINVENT is a powerful AI technology designed exclusively for de novo drug development. It provides a strong foundation for developing new compounds with drug-like characteristics. Its fundamental feature is based on rapidly exploring enormous chemical spaces in order to identify compounds with required biological activity and drug-like qualities.

The method begins with users specifying their initial constraints and desired properties for the molecules. These can include factors like diversity filters, chemical properties, and most crucially, the scoring function. These parameters are then passed to REINVENT via a JSON configuration file.

Once set up, REINVENT scores molecules using a predetermined scoring or feedback algorithm. Its purpose is to continuously optimize and suggest molecules with the highest scores based on the established criteria.

REINVENT refines these molecules in sequential steps, using a scaffold memory system to improve and stored best SMILES encountered so far. Finally, the output is delivered as SMILES, where number of SMILES is roughly equal to number of optimization steps x batch size

Slide 12: The three feedback mechanisms (10 seconds)
I have mentioned the feedback models previously, so now, I would continue to describe these models in detail.

Slide 13: Model architecture (80 seconds)

In this project, the three feedback models are integrated during the scoring stage of the REINVENT platform. These models are the scoring feedback (and I call it the baseline model), the Bradley-Terry model for pairwise comparison feedback, and the ListNet model for ranking feedback.

Despite their differences in feedback delivery, all three models share a common neural network architecture designed specifically to learn the DRD2 affinity. This common feature ensures that each model can all have the ability to learn molecules based on the DRD2 receptor's binding affinity.

The architecture uses Morgan Fingerprints, which are 2048-count bit vectors representing the molecular structure, as input to a Pytorch-based neural network. The output layer of the network features a sigmoid activation function. This setup ensures that the output, representing the predicted DRD2 affinity, remains within a range of 0 to 1.

By standardizing the architecture across different feedback mechanisms, I can all use them to classify molecules later whether they have DRD2 binding receptor.

Slide 14: Scoring (baseline) model (120 seconds)

The scoring component is a fundamental component of the REINVENT platform, designed to predict the DRD2 activity of molecules represented by SMILES strings. It uses Morgan fingerprints, as I have mentioned before.

For the model Input, Each molecule is represented by a vector of Morgan fingerprints of a single SMILES string. This representation captures the essential chemical properties of the molecule, which are crucial for predicting its biological activity.

Model Output: The output of this model is a continuous probability score between 0 and 1, indicating the likelihood of the molecule interacting effectively with the DRD2 receptor. This score is derived directly from the neural network, specifically designed to handle the complexities of molecular data.

Feedback Types:

Individual Feedback: For each SMILES string processed, the model outputs a score which serves as immediate feedback, indicating the DRD2 activity potential.
Batch Feedback: When processing multiple SMILES at once, the model outputs an array of scores. This batch feedback helps in evaluating and refining the model based on a broader dataset.
Training the Model: To train this neural network, I use Binary Cross Entropy Loss (BCELoss), which is well-suited for binary classification tasks. This loss function helps in optimizing the model by minimizing the difference between predicted probabilities and the actual labels, where ‘1’ indicates high DRD2 activity and ‘0’ signifies low activity.

By implementing such a model, I can systematically score and evaluate numerous potential drug molecules, streamlining the identification of promising candidates for further development."

Slide 15: Pairwise comparing (Bradley-Terry) model (120 seconds)

"The Bradley-Terry model is utilized in this project to facilitate pairwise comparisons between molecular candidates. This approach is particularly beneficial as it simplifies the task of determining DRD2 affinity among molecules for human chemists, who may find direct affinity assessment challenging.

Motivation: Humans often find it easier to judge between pairs of items rather than assigning absolute scores. By comparing two molecules at a time, this model allows for more intuitive feedback regarding which molecule is likely to have higher DRD2 affinity.

Model Mechanics: The model operates by defining beta_i and beta_j

as the affinities of two different SMILES strings, SMILES 1 and SMILES 2, respectively. The difference, \beta_i - \(\beta_j, dictates the outcome of the comparison. The result, pij,, is the probability derived from the sigmoid function, indicating the likelihood that SMILES 1 is more favorable than SMILES 2 in terms of DRD2 affinity.

Model Input and Output:

Input: Morgan fingerprints of two different SMILES strings.
Output: A probability score indicating if the first SMILES string has a higher predicted DRD2 affinity than the second.
Feedback Mechanism:

Individual Feedback: Each comparison results in a binary feedback—1 if SMILES 1 is judged better than SMILES 2, otherwise 0.
Batch Feedback: Involves aggregating individual feedback across multiple comparisons within a batch of SMILES, enhancing the reliability of insights derived from the model.
Training the Model: The model is trained using Binary Cross Entropy Loss (BCELoss), optimizing it to accurately predict pairwise preferences and thereby, indirectly, DRD2 affinity. The label for training is set to 1 when SMILES 1 is better and 0 otherwise, with a decision threshold of 0.5.

By leveraging this model, I can effectively harness less direct but highly valuable comparative feedback to refine the molecular design process."

Slide 16: Ranking (ListNet) model (120 seconds)

"The Ranking model used in this project, known as ListNet, employs a listwise approach to learning to rank, optimizing the order of a list of items, in this case, molecular structures, rather than just comparing individual pairs. This model is essential for situations where a more

nuanced, gradational understanding of molecule rankings is required, unlike the more binary output of pairwise models.

Motivation: While pairwise models offer valuable insights, they can sometimes be limiting by offering only win or lose feedback. The ListNet model, however, allows for a broader spectrum of feedback by providing a ranked order of molecules, offering a more refined evaluation that captures the subtleties between different molecular interactions with the DRD2 receptor.

Model Mechanics:

Input: Morgan fingerprints of three different SMILES strings.
Output: The model outputs a set of softmax scores based on the DRD2 affinity predictions from the neural network. These scores are then used to rank the SMILES strings from highest to lowest affinity.
Feedback Process:

Individual Feedback: Each SMILES string receives a rank based on its score, with the highest score indicating the strongest DRD2 affinity.
Batch Feedback: Involves generating all possible combinations of three SMILES from the output, calculating the preference scores, and then normalizing these scores to ensure a fair comparison.
Training the Model: To optimize the ranking accuracy, I train the model using the Kullback-Leibler divergence loss function. This loss measures the divergence between the predicted probability distribution and the actual, or true, softmax scores of DRD2 affinity. By minimizing this divergence, the model learns to accurately predict the relative rankings of molecular structures.

By applying the ListNet model, I can derive a more accurate and holistic view of molecular interactions, enhancing the precision of drug design research."

Slide 17: three model comparison (75 seconds)
Scoring Model:
The first model directly calculates and returns the DRD2 affinity as a single score for each SMILES string. This straightforward method provides a clear, quantifiable measure of a molecule's potential efficacy, visualized through a simple neural network diagram.

Pairwise Comparing Model:
The second model uses the Bradley-Terry formula to compare pairs of SMILES strings. It visualizes the probability that one molecule is more effective than another by processing their DRD2 scores through a sigmoid function. This model is particularly useful for visualizing direct comparisons between two molecules, enhancing decision-making.

Ranking Model:

The third model, utilizing the ListNet approach, ranks three SMILES strings simultaneously. It applies a softmax function to derive a vector of preference scores, effectively visualizing the comparative ranking based on the DRD2 affinity scores of all three molecules.

Slide 18: HITL workflow for De-Novo Drug Design (10 seconds)
Now, I would talk about HITL workflow optimized for de-novo drug design that is used in my project

Slide 19: Human-in-the-loop workflow (150 seconds)
In this Human-in-the-Loop workflow, we start with a basic model that initially guesses DRD2 affinities with low accuracy. This model, acting as a weak oracle, improves its predictions through iterative learning from human feedback.

The workflow is structured as follows: We begin each round with an empty dataset which gradually gets populated through REINVENT cycles. Each cycle includes multiple interactions with the model, where it learns from the latest feedback to refine its scoring capabilities. By comparing new SMILES strings against the model's predictions, and updating the training dataset with this new data, the model becomes more accurate over time.

Now, let’s talk about the specific settings used in this project:

We conduct three REINVENT rounds.
Each round has five human-in-the-loop interactions.
There are 56 queries per interaction to maximize learning.
The initial training dataset starts with 100 SMILES, half with known DRD2 affinity and half without.
The batch size for REINVENT is 64, with optimization steps set at 100 per batch.
By the end of three rounds, the dataset expands significantly. After processing 5 interactions across 56 queries each, the total training set grows to 1,040 SMILES, greatly enhancing the model's performance and reliability.
This structured approach allows us to systematically refine our model based on direct and quantifiable human insights

Slide 20: Active learning (AL) (60 seconds)
"Active Learning, or AL, plays a pivotal role in this project by employing a semi-supervised approach to train our machine learning model using both labeled and unlabeled data. The process begins by collecting unlabeled SMILES molecules, which are then manually labeled by human experts, referred to as 'Oracles'. This manual labeling is essential for ensuring the accuracy and relevance of the data introduced into the training set.

Once labeled, these new instances are incorporated into the model's training routine, allowing the model to continually refine and improve its predictions of DRD2 affinity. The selection of molecules for labeling is guided by a mix of strategies. Random selection introduces diversity into the training set, uncertainty-based selection targets molecules where the model shows the

least confidence—helping to close knowledge gaps—and greedy selection focuses on those molecules that the model predicts with high DRD2 affinity, optimizing for likely beneficial outcomes.

This dynamic cycle of feedback and update proceeds until the model meets our predefined conditions for accuracy and reliability. At this point, the iterative process culminates in a finely tuned model capable of making precise predictions about new, unseen molecular structures."

Slide 21: Molecular descriptor filtering for REINVENT's generated SMILES (75 seconds)
"In my project, I utilize molecular descriptor filtering to select the most promising SMILES generated by REINVENT. These filters are based on properties that significantly impact a molecule's pharmacokinetic behavior, allowing me to characterize and assess the physical, chemical, and structural attributes essential for effective drug candidates.

To ensure that each molecule has a potential for high DRD2 affinity, I aim to maximize this particular descriptor within the range of 0.75 to 1.0, which is the score bracket where molecules have shown significant affinity. For properties like logP, molecular weight, and hydrogen bond donors and acceptors, I apply thresholds based on Lipinski's Rule of Five. This rule helps me filter out molecules likely unsuitable for drug development due to poor absorption or permeability.

Additionally, I use the Total Polar Surface Area (TPSA) and the number of rotatable bonds to further refine my selection, ensuring that each molecule is not just potent but also viable in a real-world biomedical context. These thresholds guide the inclusion or exclusion of molecules, streamlining the selection process towards those that are most likely to be bioactive and efficient as therapeutic agents."

Slide 22: Other metrics for REINVENT generated SMILES besides DRD2 affinity (75 seconds)

"In my project, I don't just focus on DRD2 affinity but also consider several other crucial metrics to evaluate the quality of molecules generated by REINVENT. One of these is the novelty score, which gauges how many of the molecules are not already present in the training set. I aim for most molecules, ideally more than 95%, to be novel, ensuring we're exploring new chemical space.

Another important metric is the diversity score, which measures the chemical variety within the generated set. A diversity score above 0.7 is preferred as it indicates a broad range of potentially effective drug candidates.

The synthetic accessibility score, or SA score, helps me determine how easily a molecule can be synthesized. Here, a lower score is better, with the target typically set below 3, making the synthesis process feasible.

Lastly, the quantitative estimate of drug-likeness, or QED score, evaluates how drug-like the compounds are based on several molecular properties. A higher QED score, close to 1, is desirable as it suggests that the molecule adheres to known pharmacological profiles needed for successful drugs."

Slide 23: Results (10 seconds)
"Let's now review the results achieved from applying these methodologies in enhancing the predictive accuracy and relevance of generated molecules."

Slide 24: ROC curve evolution after 3 REINVENT rounds (30 seconds)
In this slide, I display the ROC curve evolution across three rounds of REINVENT, using three different feedback models. Initially, the pre-trained models started with low AUC values due to limited data on DRD2 affinity. However, with each HITL iteration, there's a noticeable improvement in AUC for all models, demonstrating the effectiveness of incorporating human feedback into the learning process.

Slide 25: ROC curve comparison between 6 running cases (90 seconds)
"In this analysis, I compared the performance of all feedback models under different conditions to understand their robustness and reliability. Each graph represents a combination of three acquisition strategies—random, uncertainty, and greedy—paired with two levels of noise, 0.0 and 0.1, in the Oracle's predictions.

The ROC curves you see reflect the evolution of model accuracy over three rounds of HITL iterations. Initially, each model begins with modest AUC values since they start with limited DRD2 affinity data. As the iterations progress, significant improvements are observed due to the incorporation of human feedback.

Notice how the ranking model, displayed in green, generally performs the best across almost all scenarios, especially in noise-free environments. It's followed closely by the comparing model in blue, which also shows robust performance even as noise levels increase.

These results demonstrate the effectiveness of integrating human feedback into the learning cycle, with each model adapting and enhancing its predictive accuracy based on the specific acquisition strategy and noise resilience."

Slide 26: Classification metrics of user feedback models (90 seconds)
In evaluating the classification performance of user feedback models, I focused on multiple metrics across different configurations to identify the most effective approach in predicting DRD2 affinity. This chart shows the precision, recall, F1 score, and Matthews Correlation Coefficient (MCC) for various combinations of feedback models and acquisition strategies.

Precision is consistently high across most models, indicating that when a positive DRD2 affinity is predicted, it's likely correct. However, the real challenge lies in improving recall and MCC, especially since DRD2 active compounds are rare in the dataset. This underscores the importance of not just precision but also a balanced recall to avoid missing potential hits.

The ranking model under random acquisition with no noise stands out, demonstrating superior recall and MCC compared to others. This suggests that the ranking model, which evaluates multiple molecules together, is more adept at capturing nuances in chemist feedback compared to binary or pairwise comparisons.

Overall, comparing and ranking feedback models have shown to deliver more reliable classification results than the baseline scoring model. The data also reveals that incorporating human judgment through random sampling significantly enhances the MCC, indicating a robust model performance across varied conditions.

Slide 27: Percentage of generated SMILES that satisfy each filter at all REINVENT rounds (90 seconds)

Most of the molecules I generated using REINVENT meet the established filters, except for LogP and the number of hydrogen bond donors. These properties prove more challenging to optimize. The DRD2 affinity for the generated molecules is determined by the Oracle. Among the models, the scoring model is the most effective in generating molecules with high DRD2 affinity. It maintains high percentages across multiple rounds and conditions, indicating its robustness and reliability.

The comparing model performs well, following closely behind the scoring model. However, the ranking model, despite being robust, shows a slight lag in achieving high DRD2 affinity percentages compared to the other models. This data emphasizes the relative strengths and weaknesses of each model in generating viable molecules.

While all models are capable of producing molecules that meet the desired pharmacokinetic properties, the scoring model stands out for its efficiency and effectiveness in targeting DRD2 affinity. This capability ensures that the molecules generated are not only novel but also align well with the desired properties, highlighting the importance of model selection based on specific research goals in de novo molecular design.

Slide 28: DRD2 affinity of REINVENT's generated molecules for all rounds (75 seconds)

For generating molecules with the highest DRD2 affinity, I found that the scoring model consistently outperforms the others. It achieves higher affinity values across various acquisition strategies and noise levels, both for filtered and unfiltered SMILES. The comparing model follows closely, showing good performance but slightly less effective than the scoring model.

The ranking model, although effective, generally trails behind the scoring and comparing models in achieving high DRD2 affinity.

Unfiltered molecules generated by the scoring model typically start with higher DRD2 affinity and maintain this advantage even after filtering. The comparing model also shows significant improvement post-filtering, but not to the same extent as the scoring model. The ranking model, while improving after filtering, still lags behind the other two.

Across different acquisition strategies, the scoring model's performance remains robust, making it the most reliable for generating molecules with high DRD2 affinity. The results indicate that leveraging the scoring model is most advantageous when the goal is to maximize DRD2 affinity in the generated molecules.

Slide 29: Metric scores comparison between all running cases for filtered SMILES (75 seconds)

The comparison of metric scores for filtered SMILES across different models reveals some key insights. The novelty score is nearly equal for all three models, with the comparing model frequently achieving the highest scores. When evaluating diversity scores, I see that the ranking model consistently outperforms the comparing model, which in turn surpasses the scoring model.

Interestingly, the synthetic accessibility (SA) score appears to be inversely proportional to the diversity score. Consequently, the scoring model exhibits the best SA score, followed by the comparing model, with the ranking model trailing behind. Lastly, the QED score, which measures drug-likeness, is consistently higher for the ranking model compared to the comparing model and the scoring model. This consistent performance of the ranking model in QED suggests its effectiveness in producing drug-like molecules, although it may not always excel in other areas like synthetic accessibility. Overall, these insights highlight the strengths and weaknesses of each model in generating high-quality molecules based on various important metrics.

Slide 30: Example of best molecules for each running case (30 seconds)
Molecules generated by the scoring model tend to have simple structures with fewer branches and fewer functional groups. In contrast, the comparing and ranking models produce molecules with greater structural complexity and diversity, incorporating different ring systems and functional groups. These more complex structures manage to balance drug-likeness, diversity, and novelty more effectively than the simpler molecules from the scoring model.

Slide 31: Discussions (10 seconds)

Let's move on to the discussions of the previous results I have shown

Slide 32: Comparison of feedback models (75 seconds)

In comparing the feedback models, I can see distinct advantages and disadvantages for each. The scoring model performs the worst in learning chemist preferences but excels in generating molecules that maximize the average probability of DRD2 affinity and SA score. On the other hand, the ranking model is superior in learning chemist preferences, achieving high MCC scores for classification, passing most molecular descriptors, and maximizing molecular diversity and QED scores. The comparing model, also known as the Bradley-Terry model, strikes a balance between the scoring and ranking models. It performs moderately well across all criteria, particularly excelling in generating molecules with high novelty scores. This balance allows the comparing model to be versatile, though it may not lead in any single category. The comparative analysis helps me understand how each model aligns with different objectives in molecular generation and preference learning.

Slide 33: Comparison of acquisition strategies (75 seconds)
In comparing the acquisition strategies, I observe that random feedback acquisition leads to the highest molecular novelty and diversity. However, it produces molecules that often do not pass most molecular descriptor filters and the DRD2 affinity threshold. Greedy feedback acquisition excels in generating molecules that maximize DRD2 affinity and minimize SA scores but performs the worst in terms of novelty and diversity. Uncertainty acquisition balances performance, consistently placing second in most categories. This balance makes it a reliable strategy across various metrics. By understanding the strengths and weaknesses of each acquisition strategy, I can better align the model's performance with specific project objectives, whether it be enhancing novelty, optimizing affinity, or balancing multiple molecular properties.

Slide 34: Comparison of labeling noise (75 seconds)
Non-noisy feedback models, where the human acts as an oracle, excel in producing molecules that maximize DRD2 affinity and improve synthetic accessibility. These models perform better in terms of DRD2 affinity, ensuring that the generated molecules are more likely to be potent and synthetically accessible. On the other hand, models with added noise of 0.1 outperform in generating molecules that pass molecular descriptor filters, while also maximizing novelty, diversity, and QED scores. This added noise seems to aid in generating a broader variety of molecules, enhancing chemical diversity and drug-likeness. However, it remains unclear how the noise impacts the learning performance regarding chemist preferences. This uncertainty suggests that while noise can help in generating diverse and drug-like molecules, its effect on accurately capturing and replicating expert chemist knowledge and preferences needs further investigation.

Slide 35: Conclusions (10 seconds)

That was everything about my research project, and now we arrive at some conclusions that I believe can be inferred from my observations of the results.

Slide 36: Conclusion content (90 seconds)

The feedback mechanism, acquisition function, and noise in labeling offer robust customization in de-novo molecular design. Users can change these to fit specific molecule generation objectives. For generating molecules that target chemical properties like DRD2 affinity and synthesizability, I find that using the scoring model with a greedy acquisition strategy and no labeling noise works well. If the focus is on generating novel molecules while maintaining a balance across all metrics, the pairwise comparing Bradley-Terry model with an uncertainty acquisition strategy is effective. For those aiming to generate diverse and drug-like molecules, the ranking ListNet model with a random acquisition strategy and labeling noise is the best choice.

Future research could refine these feedback models or introduce new mechanisms, such as multiple binary preferences for molecular properties or molecule editing. This approach proposes modified molecular structures that might be more promising than the original ones. I can validate these findings through real human experiments where medicinal chemists interact with the feedback system via a graphical user interface.

==========

Possible QA:

What motivated you to undertake this project? My motivation for this project is to enhance human-in-the-loop (HITL) assisted de novo molecular design. While this framework already exists, my contribution is the introduction of different modes of human interaction, such as pairwise comparison and ranking, beyond the traditional binary like/dislike feedback. This approach aims to determine which user feedback mechanism is most effective in learning chemist preferences about molecules, thereby improving drug design targeting specific physical properties and achieving favorable molecular descriptors.

Can you describe the three feedback models you implemented? I developed three feedback models for the REINVENT platform: the scoring model, the pairwise comparing (Bradley-Terry) model, and the ranking (ListNet) model. The scoring model predicts the probability of DRD2 activity for a single SMILES string. The pairwise comparing model uses the Bradley-Terry formulation to compare pairs of SMILES strings, determining which is more likely to have DRD2 activity. The ranking model, using the ListNet architecture, ranks a list of SMILES strings based on their predicted DRD2 activity, aiming to optimize the ranking directly. Each model employs a neural network to learn DRD2 affinity, providing diverse feedback mechanisms.

Why did you choose to explore different feedback mechanisms? I chose to explore different feedback mechanisms to push the boundaries of the state-of-the-art in HITL-assisted molecular design. By introducing novel feedback methods like pairwise comparisons and ranking, I aim to identify which type of user feedback is most efficient in learning chemist preferences and improving molecular generation. This approach provides recommendations for

generative chemistry users on which feedback mechanism works best to enhance specific criteria, such as targeting physical properties or achieving favorable molecular descriptors.

How do these feedback mechanisms improve the HITL process in molecular design? These feedback mechanisms improve the HITL process by offering more nuanced and effective ways for chemists to interact with the generative model. Instead of relying solely on binary feedback, the pairwise comparing and ranking methods allow for more detailed preferences to be captured, leading to better learning of chemist preferences. This results in the generation of molecules that are more likely to meet specific design criteria, ultimately enhancing the efficiency and effectiveness of the drug design process.

What are the practical applications of your findings? The practical applications of my findings include providing a robust framework for medicinal chemists to generate molecules that meet targeted physical and chemical properties. By utilizing the most effective feedback mechanisms identified in this project, chemists can improve the accuracy and relevance of their molecular designs, leading to the development of more efficient and effective drugs. This approach can be extended to other aspects of drug design, such as optimizing molecular descriptors and improving drug-likeness.

How did you validate the performance of the different feedback models? I validated the performance of the different feedback models by evaluating their ability to learn chemist preferences and generate molecules with desired properties. This involved comparing the ROC AUC values, precision, recall, F1 scores, and Matthews correlation coefficient (MCC) for each model across multiple HITL iterations. Additionally, I assessed the generated molecules' compliance with molecular descriptor filters and their DRD2 affinity to ensure practical relevance in drug design.

What challenges did you face during the implementation of the feedback models, and how did you overcome them? One of the main challenges was integrating the new feedback mechanisms into the existing REINVENT platform. To overcome this, I carefully designed the neural network architectures to handle different types of feedback and optimized the training process for each model. Another challenge was ensuring the feedback accurately reflected chemist preferences, which I addressed by iterating on the feedback collection and processing methods, incorporating noise handling, and validating results through multiple experiments.

How do the different acquisition strategies (random, uncertainty, greedy) affect the performance of the feedback models? The different acquisition strategies have distinct impacts on model performance. The random acquisition strategy introduces diversity and novelty in the generated molecules but may not always produce the most promising candidates. The uncertainty acquisition strategy focuses on molecules where the model has low confidence, which helps improve the model's robustness. The greedy acquisition strategy targets molecules with the highest predicted DRD2 affinity, leading to high-quality candidates but potentially lower diversity. By analyzing these strategies, I found that each has its strengths and weaknesses, and their effectiveness can vary depending on the specific goals of the molecular design process.

What role does labeling noise play in the feedback mechanisms, and how did you address it? Labeling noise can significantly impact the performance of feedback models by introducing variability in the feedback provided by chemists. I addressed this by experimenting

with different levels of noise (0.0 and 0.1) and observing their effects on model performance. I found that non-noisy feedback models generally performed better in producing molecules with high DRD2 affinity and improving synthetic accessibility. Noisy feedback models, however, excelled in producing molecules that passed molecular descriptor filters and maximized novelty, diversity, and QED scores. This analysis helped identify the optimal balance between noise and feedback accuracy for different design objectives.

What future improvements would you suggest for the feedback models and the overall HITL framework? Future improvements could include refining the feedback models by incorporating more complex user feedback mechanisms, such as multiple binary preferences for different molecular properties or detailed feedback on specific structural features. Additionally, introducing real-time interactive feedback through a graphical user interface (GUI) could enhance chemist engagement and accuracy in providing feedback. Further research could explore the integration of advanced machine learning techniques, such as reinforcement learning, to dynamically adapt the feedback mechanisms based on evolving chemist preferences and design goals.

How did you ensure that the generated molecules are chemically valid and synthesizable? Ensuring chemical validity and synthesizability involved applying molecular descriptor filters such as Lipinski's Rule of Five, Veber et al.'s rules, and other known drug-likeness rules. These filters helped eliminate molecules that were unlikely to be effective drugs. Additionally, I utilized synthetic accessibility (SA) scores to assess how easily a molecule could be synthesized. By filtering out molecules with low SA scores, I ensured that the generated molecules were not only theoretically effective but also practically feasible to produce.

Can you explain the significance of using DRD2 affinity in your project? DRD2 affinity is significant in this project because the dopamine D2 receptor (DRD2) is a crucial target in the development of drugs for neurological and psychiatric disorders such as schizophrenia and Parkinson's disease. High DRD2 affinity indicates that a molecule is likely to interact effectively with this receptor, making it a promising candidate for therapeutic development. By focusing on DRD2 affinity, the project aims to generate molecules with potential clinical relevance, addressing unmet medical needs.

What criteria did you use to select the initial training dataset, and how did it influence the results? The initial training dataset consisted of 100 SMILES with known DRD2 activity and 100 without DRD2 activity. This balanced dataset provided a solid foundation for training the feedback models, allowing them to distinguish between active and inactive molecules effectively. The choice of this dataset influenced the results by ensuring that the models had enough information to learn meaningful patterns and make accurate predictions from the start. It also helped in evaluating the performance improvements after each HITL iteration.

How does the Human-in-the-Loop (HITL) approach improve over traditional machine learning methods in molecular design? The HITL approach improves over traditional machine learning methods by incorporating real-time feedback from human experts, allowing the model to learn and adapt based on expert knowledge and preferences. This interactive process helps refine the model iteratively, leading to more accurate and relevant outcomes. In molecular design, HITL ensures that generated molecules align better with the desired

properties and therapeutic goals, as human feedback can provide nuanced insights that purely data-driven methods might miss.

What are the practical implications of your findings for the pharmaceutical industry? The practical implications of my findings for the pharmaceutical industry include providing a more efficient and targeted approach to drug discovery. By identifying the most effective feedback mechanisms and acquisition strategies, the project offers a framework for optimizing the design of novel compounds. This can lead to faster development of drug candidates with desired properties, reducing time and costs associated with the early stages of drug discovery. Furthermore, the ability to customize feedback mechanisms based on specific goals enhances the versatility and applicability of the HITL framework in various therapeutic areas.

How did you handle the potential bias in human feedback during the HITL process? To handle potential bias in human feedback, I incorporated multiple rounds of feedback from different experts to ensure a diverse set of opinions and minimize individual biases. Additionally, the feedback was averaged and normalized to reduce the influence of outliers. I also employed a noise model to simulate human errors and variations, which helped the system to generalize better and become more robust against biased feedback.

What were the main challenges you faced while implementing the feedback models, and how did you overcome them? One of the main challenges was ensuring the accurate integration of human feedback into the machine learning models. I overcame this by developing a robust data processing pipeline that could handle and standardize the feedback efficiently. Another challenge was maintaining a balance between model complexity and computational efficiency. To address this, I performed extensive hyperparameter tuning and used efficient algorithms to optimize performance without compromising accuracy.

Can you elaborate on how the ranking model differs from the scoring and comparing models in terms of performance? The ranking model differs from the scoring and comparing models by directly optimizing the ranking of a list of molecules based on their predicted DRD2 affinity scores. This model uses the ListNet architecture, which focuses on the relative order of molecules rather than their absolute scores. In terms of performance, the ranking model consistently achieved higher MCC and QED scores, indicating its effectiveness in capturing human preferences and generating diverse, drug-like molecules. However, it may not always produce molecules with the highest individual DRD2 affinity scores compared to the scoring model.

What steps did you take to validate the performance of your feedback models? I validated the performance of the feedback models through cross-validation on a separate test dataset comprising 1000 SMILES with known DRD2 activity and inactivity. I evaluated various metrics such as ROC AUC, MCC, precision, recall, F1 score, and QED score to ensure comprehensive assessment. Additionally, I compared the performance across different acquisition strategies and noise levels to understand the robustness and generalizability of the models. Real-world validation involved collaborating with chemists to provide practical feedback on the generated molecules.