ML4CC: Lecture 12

Sit with your discussion groups (same as last time)

Assignments reminder

No more PMIRO+Q!

Today - Project check-ins **during second half of class**

Apr 24 - Exam II (during first half; second half of class is project work time)

May 1 - Project Presentations

Project reports due May 8th.

Summary of last paper

P - Solving the OPF is too computationally expensive. Let’s find a way to reduce the number of equations included in it

M - Train a graph neural network to predict which lines will be congested and only include those line constraints in the OPF

I - Use of GNNs to create a reduced OPF

R - Runtime of the ROPF is less than the full OPF, but there are errors that would lead to over-capacity lines

O - How would this scale to more realistic grids?

Climate Change in the News

CLIMATE TECH

Exclusive: Occidental Petroleum Buys DAC Startup Holocene

That makes two direct air capture acquisitions for the oil and gas major.

EMILY PONTECORVO · APRIL 16, 2025

Like every other climate tech industry, direct air capture startups have faced a great deal of uncertainty since Trump took office. While the technology has historically had bipartisan support, the Trump administration has been excising programs and projects with seemingly any connection to climate change. It has hollowed out the Department of Energy's carbon dioxide removal team, my colleague Katie Brigham reported in February, leaving just one employee overseeing the $3.5 billion Direct Air Capture Hubs program that was authorized by the Infrastructure Investment and Jobs Act. Additional cuts at the Office of Clean Energy Demonstrations, which also has a role in overseeing the program, or even a potential closure of that office, are expected in the coming weeks. The Direct Air Capture Hubs were also on a list of grants the administration was considering trying to cut.

Holocene's technology is similar to that of Carbon Engineering. Both companies use fans to pull air into a closed system, where it passes through a liquid with a unique chemistry that attracts CO2. In the case of Carbon Engineering, the carbon in the air binds with potassium hydroxide in water; in Holocene's system, it binds with amino acids. Then both companies react that carbon-rich water with another chemical that further concentrates the CO2 into solids that can be filtered out. The last step is heating those solids, releasing the CO2 so that it can be sequestered underground.

Holocene's advantage — and the reason it thinks it can achieve $100 per ton carbon removal — is that it uses a unique chemistry that requires relatively low heat to separate the CO2. Whereas Carbon Engineering uses natural gas for that final step, Holocene told me it could use renewable electricity, or even waste heat from a data center.

Oxy's acquisition of Carbon Engineering was controversial among climate advocates. While many see direct air capture as a promising way to clean up the excess carbon that will remain in the atmosphere even after emissions decline, skeptics worry that oil companies will use it as justification to keep producing oil — a fear that Oxy has not exactly allayed.

The company plans to take some of the carbon it captures and sequester it in dedicated carbon storage wells. It signed a deal to sequester 500,000 tons of carbon on behalf of Microsoft last year. But it will also pump carbon into aging oil wells to increase oil production, a process called enhanced oil recovery. In the past, Oxy's CEO Vicki Hollub has framed its investments in direct air capture tech as a way to produce "net-zero oil," and as a "license to continue to operate" as an oil producer.

Paper 10 Discussion

AI assisted Search for Atmospheric CO2 Capture

Shivshankar

Abstract

Carbon capture technologies is an important tool for mitigating climate change [42]. In recent years, polymer membrane separation methods have emerged as a promising technology for separating CO2 and other green house gases from the atmosphere. Designing new polymers for such tasks is quite difficult. In this work we look at machine learning based methods to search for new polymer designs optimized for CO2 separation. An ensemble ML models is trained on a large database of molecules to predict permeabilities of CO2/N2 and CO2/O2 pairs. We then use search based optimization to discover new polymers that surpass existing polymer designs. Simulations are then done to verify the predicted performance of the new designs. Overall result suggests that ML based search can be used to discover new polymers optimized for carbon capture.

Attendance

Select one person from the group to be the attendance taker and fill out the attendance form (linked to under syllabus in Brightspace)

Discussion Question 1

What is the Robeson upper bound and how does it relate to a Pareto front (in your own words)?

Description of tradeoff between permeability and selectivity

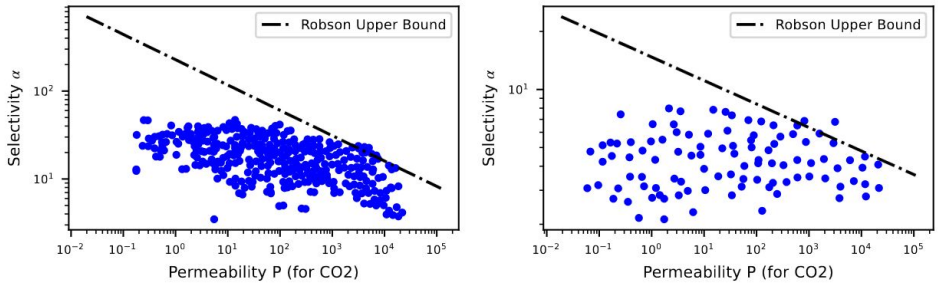


Figure 1: Double logarithmic plots of selectivity versus CO2 permeability for gas pairs a) CO2/N2 and b) CO2/O2 from a database of existing molecules (e.g. PolyInfo[32],PI1M[27]). The Robeson upper-bound is marked in black line

The Robeson upper bound describes the observed negative relationship between permeability and selectivity. Most known polymers do not have better permeability/selectivity value pairs than this upper bound.

In this way, the upper bound represents the non-dominated front, i.e. the Pareto front, for the multi-objective function problem of trying to maximize both permeability and selectivity

Discussion Question 2

What part of this work could be considered an “emulator” (based on how we’ve previously used that word in class)?

Emulator: an ML system that circumvents running expensive simulations

ML-aided Polymer Design The accurate and efficient property prediction is essential to the design of polymers in various applications[26]. Traditionally, chemists and polymer scientists have relied on physical simulation to study and predict properties of various molecules. These simulations while accurate, are compute-intensive. On the other hand deep-ML models provide a quick inference method, which has led to them their rising popularity in the field of property prediction [14, 1, 22, 10, 29, 30]. This has been further aided by the availability of multiple large scale datasets [39, 48, 32, 49], that are needed to develop accurate deep-learning models [41, 43, 18, 28]. Recently, with the advent of large language models, large transformer inspired architectures for doing property predictions are also becoming popular [18, 45, 20].

The first step is to train machine learning (ML) models to predict the required properties, such as selectivity and permeability, using existing training sets like PoLyInfo[32] and PI1M [27]. We refer to this model as a Property Prediction Function (PPF) in our work. Once a well-trained and calibrated

The PPF is an emulator that predicts permeability and selectivity without needing to run a simulation (note: this is “multi-task” learning)

Discussion Question 3

What are the two instances where an ensemble method is used here?

Ensembles: combining outputs of models trained separately

Features and ML Model We train an ensemble model consisting of various machine learning algorithms, such as elastic nets, boosted random forests (Xgboost), feed-forward neural networks (FFN), graph convolutional networks (pGCN [33]) and Transformers (TransPolymer [45]) For the

GCNs, were trained using both unsupervised losses and multi-task regression losses. To determine the best ensemble, we conducted a grid search to optimize the hyperparameters and chose the best ensemble based on cross-validation results.

This is itself an ensemble model!

Discussion Question 4

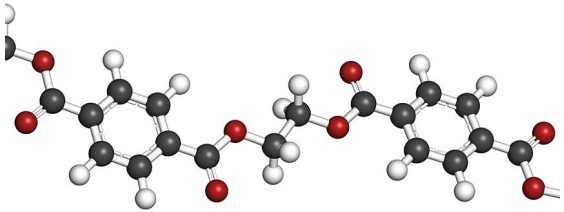
How do the inputs to the elastic net and XGBoost models differ from the inputs to the Graph Neural Network model or the inputs to the Transformer model?

Hand-designed features for simpler models

(FFN), graph convolutional networks (pGCN [33]) and Transformers (TransPolymer [45]) For the classical models, we used chemically relevant descriptors including bond angles, ring type and substructures, RDKit chemical descriptors [21] and Morgan fingerprints [38], among others. These features are widely used in the literature and are known to be important for predicting various properties of molecules [6]. In addition, we used learned embeddings for substructures and important atoms in the neural network models. The larger deep-learning models, such as Transformers and GCNs, were trained using both unsupervised losses and multi-task regression losses. To determine

Full molecule data for complex models

Graph Neural Networks for molecules



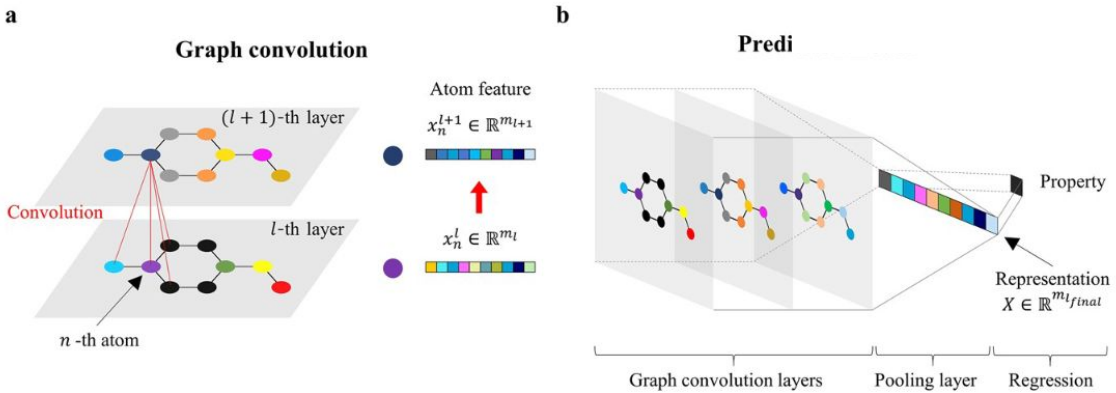
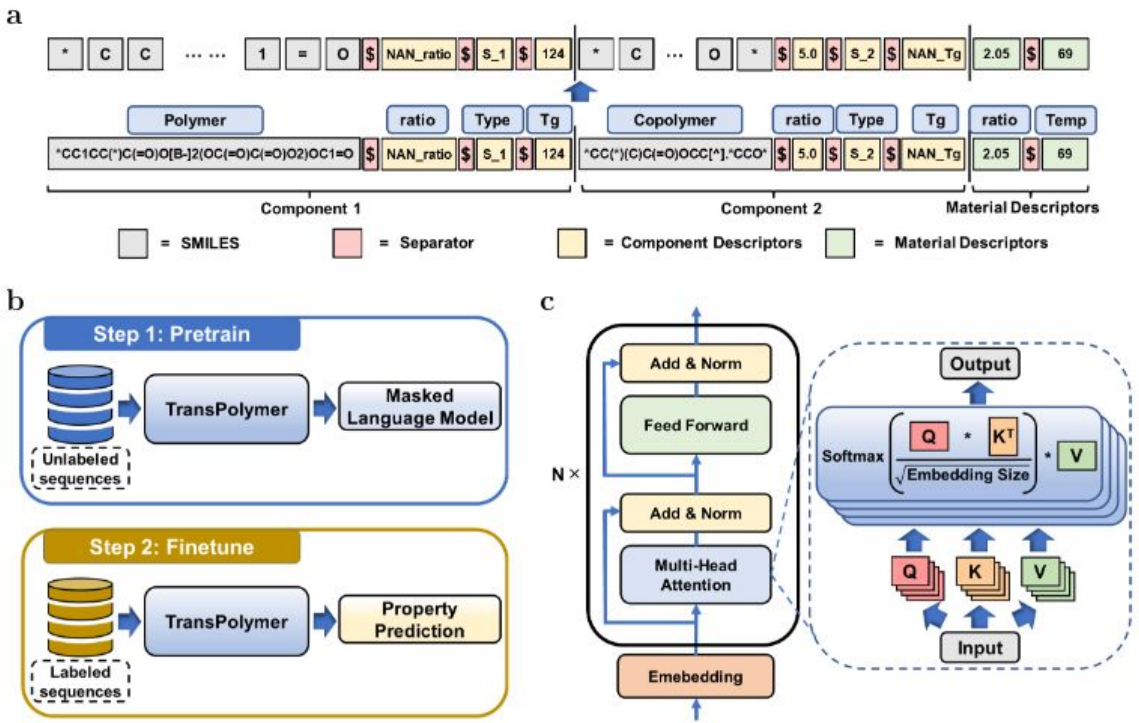


Figure 2. Property prediction model based on the GCN. (a) Feature vector of n-th atom x_n^l is updated iteratively through the l-th convolutional layer by the graph convolution that aggregates the features of neighboring atoms. (b) Graph-level molecular representation vector X is obtained by the pooling layer that sums up the feature vectors of all the atoms at the final convolution layer. The vector X is input to the regression algorithms (linear regression (LR) only shown here) for the property prediction.

Park et al, 2022

Full molecule data for complex models

Transformers for molecules: need to represent the molecular structure as a sequence of characters



Xu et al, 2022

Discussion Question 5

What are the observations/states, actions, and rewards for the reinforcement learning problem here? What type of model is used to generate the policy?

Making better polymers

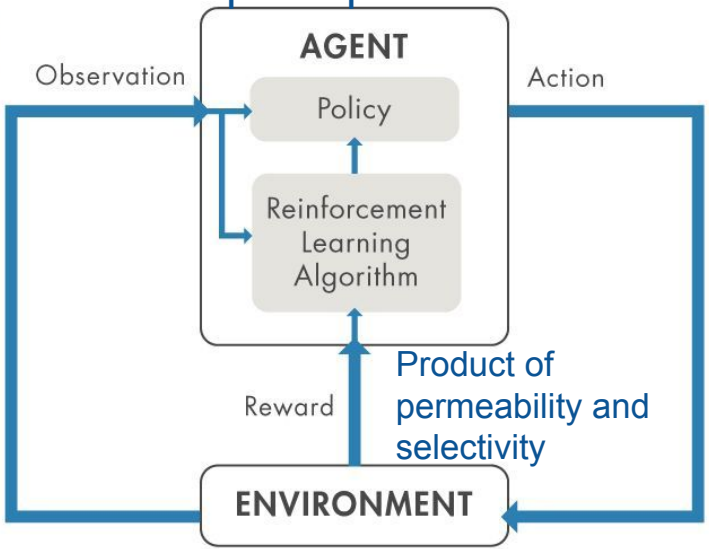
Environment and Score/Reward Function The score function is the reward provided for the agent to maximize. Ideally we want a polymer which has high CO2 permeability while also high selectivity for CO2 compared to O2 and N2. A natural metric is the product of the pair selectivities and CO2 permeability [48]. We use the estimate of selectivity and permeability produced by the PPF.

State and Policy Network Each state vector, represents a specific intermediate molecular configuration. The policy network takes the current graph/molecule as state and chooses an action which is applied to the current molecule to modify it. Each policy action is specified via three steps: I), the identification of the specific site where a new fragment is to be introduced; II, the selection of the new fragment itself; and III), the precise bonding site on the new fragment.

We use three different multi-layered perceptrons M_1, M_2, M_3 to parameterize each individual decision of the three step policy action. Following the approach of Hwang *et al.* [17], Yang *et al.* [47], we represent the current state molecule as an undirected graph, using a graph convolutional network to obtain node embeddings that are then aggregated to produce a graph embedding. In the

Set of 3 multi-layer perceptrons

Structure of the polymer passed through a Graph Neural Network

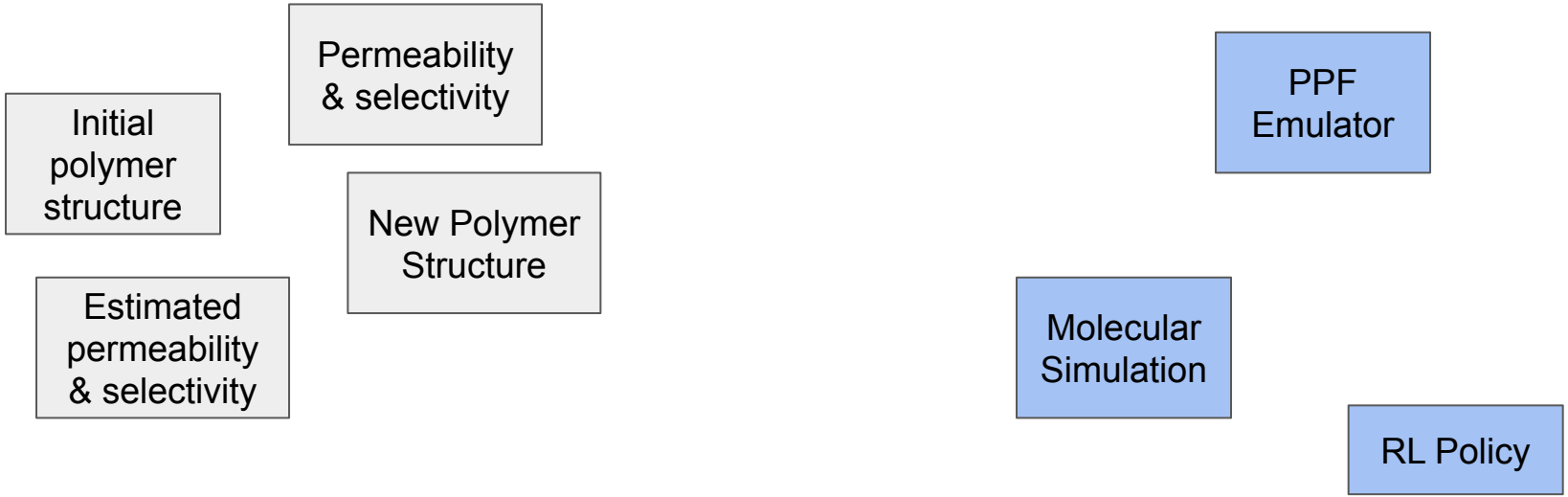


Product of permeability and selectivity

Change a molecule on the polymer

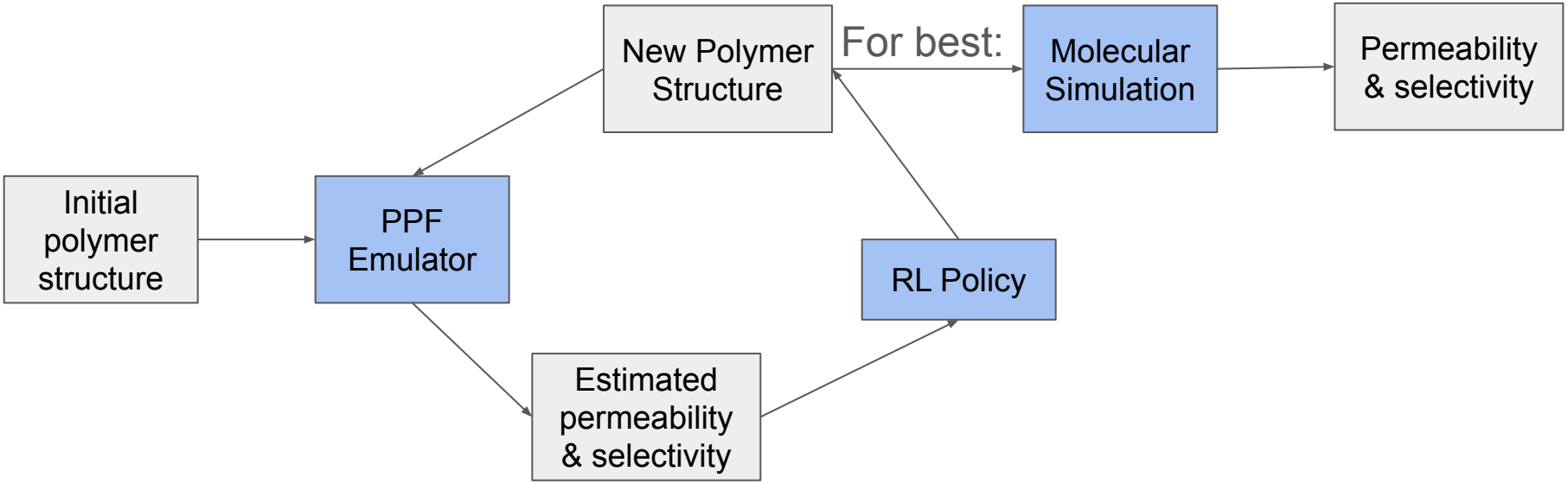
Discussion Question 6

Create a flow chart of the methodology based on these parts (hint: there is a loop):



Discussion Question 6

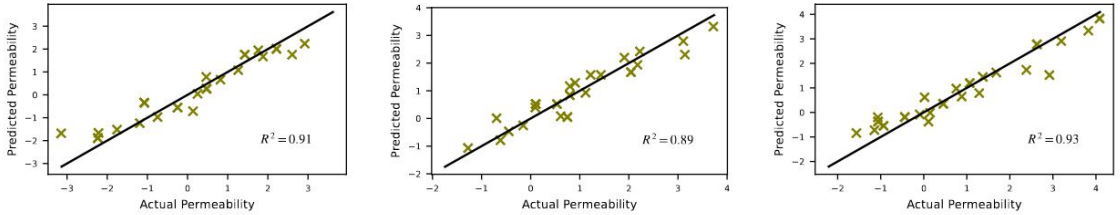
Create a flow chart of the methodology based on these parts (hint: there is a loop):



Discussion Question 7

For the polymers discovered by the model, does the PPF output align with the results of physical simulation (hint: check appendix)? Why was it important to check? Do these molecules perform better than the Robeson bound?

Good alignment between PPF and molecular simulation

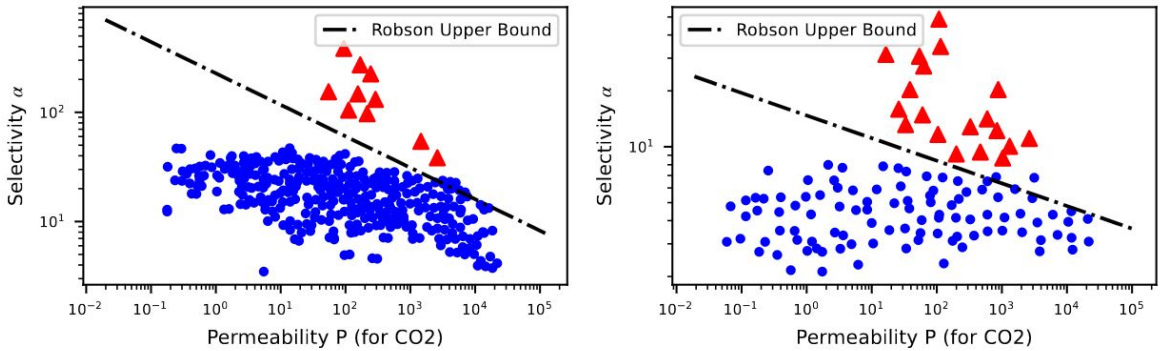


(a) N2 (b) O2 (c) CO2

Figure 4: Comparison of the prediction performance of the ensemble ML model trained to predict permeabilities for a) N2 , b) O2 and c) CO2 on generated polymers where we take molecular simulation results as the actual permeability

Important to validate the PPF guesses, because the RL policy may have found some really weird solutions that are outside the space of polymers the PPF was trained on.

Designed polymers perform better



(a) CO2/N2 (b) CO2/O2

Figure 2: Double log plots, overlaying experimental data from existing databases (blue points), the Robeson bound (black) and a sample of novel molecules found by our method (red). The figures show results for the gas pairs a) CO2/N2 and b) CO2/O2. It is clear that our method generates molecules that surpass both upper bounds

But:

Limitations Our framework while easily adjustable to incorporate other important metrics like synthetic accessibility, does not take them into consideration. Even ignoring questions of accessibility, there are other factors important for gas separation that are ignored. A polymer's gas filtration performance also depends heavily on factors like morphology of the polymer, operating temperature etc [34] which have been ignored here. Moreover, we have focused on separating CO2 from N2

Discussion Question 8

Share what questions you wrote in your PMIRO+Q and decide as a group what you'd like to ask.

Update your PMIRO+Q

Submit a second file to the Brightspace assignment (don’t overwrite the original):

It should:

Update your PMIRO as needed

Answer your own Q

You can be talking with your group during this!

15 min break

Project Presentations

8 minutes + 1 minute for questions (staying on time and asking questions will be part of the grade)

Everyone must be on time to class as we will need to start immediately

Order will be chosen randomly

You will need to upload your slides to Google Drive before the start of class (the title should include your team name)

Exam Prep

Exam 2 has the same format as Exam 1 (you will be asked about climate content, ML ideas, and the papers).

It will cover content related to the last five papers.

As before, focus on knowing the material covered in the slides.

Additional notes

You don’t need to know:

- How DQN (or any other specific RL algorithm) works.
- Specifics on how energy is generated by different methods
- Details of how different methods pull CO2 from the air

You do need to know methods used in paper 10 that were introduced in the first half of the course

Project Check-Ins

Each team must come up to me at some point before the end of class to discuss progress (don’t wait til the end!!)