Towards Analyzing and Minimizing the Carbon Footprint of Reinforcement Learning-Based Sequential Recommender Systems

Alejandro Plaza & Denis Parra
Pontificia Universidad Católica de Chile

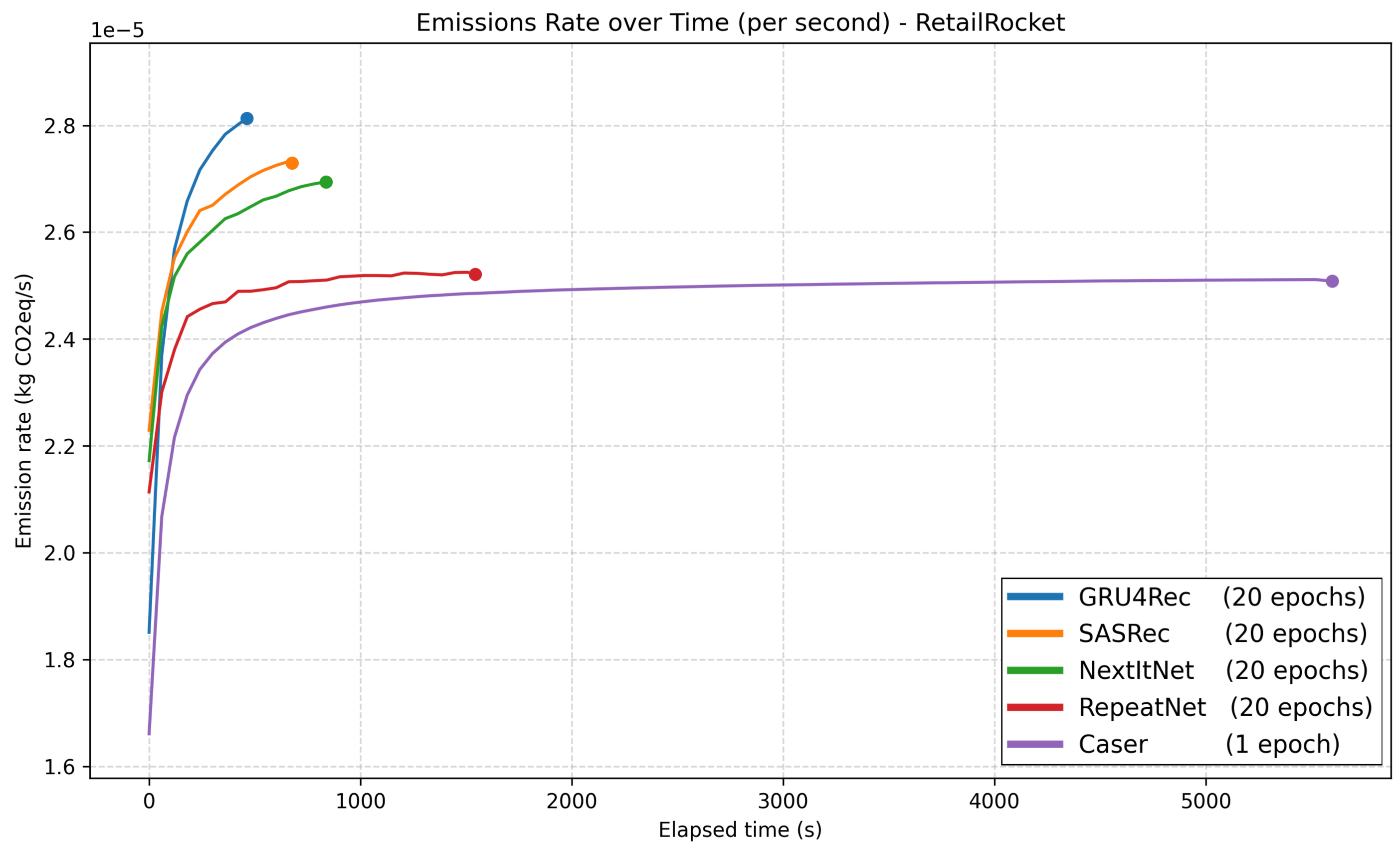
Introduction

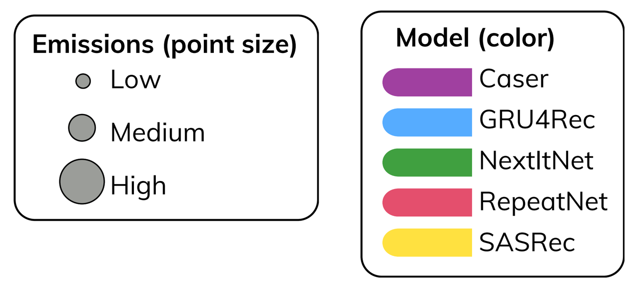
Recommender systems have become central to digital platforms, including e-commerce and streaming services, but their computational demands continue to grow (Beel et al., 2024). Recent work shows that **training modern deep learning recommenders can emit significant amounts of carbon equivalent (CO2e)**. Although prior research has explored emissions in collaborative filtering (Wegmeth et al., 2025), only a few on sequential models and, to the best of our knowledge, no work has analyzed reinforcement learning–based recommenders, creating a gap we aim to address.

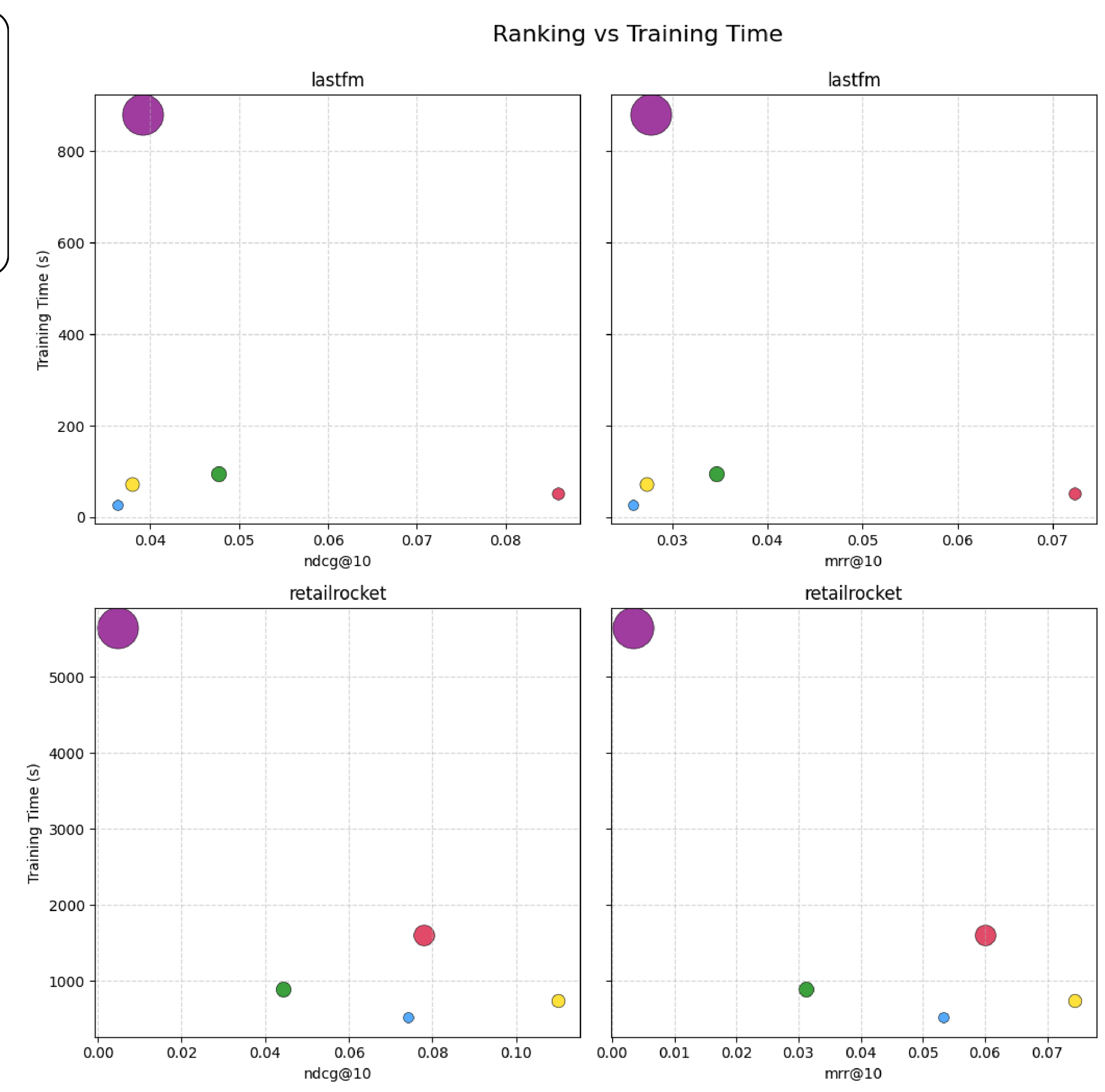
We hypothesize that reinforcement learning-based sequential recommender systems can achieve competitive performance without increasing their carbon footprint, provided that model configurations and hyperparameters are carefully selected.

Sequential Recommenders Experiments

In the first stage, we compare the carbon footprint and recommendation quality of five sequential recommenders. All measurements are collected using a modified version of the CodeCarbon library, enabling tracking of cumulative emissions over time.







* Models trained for 20 epochs, except Caser (1 epoch).

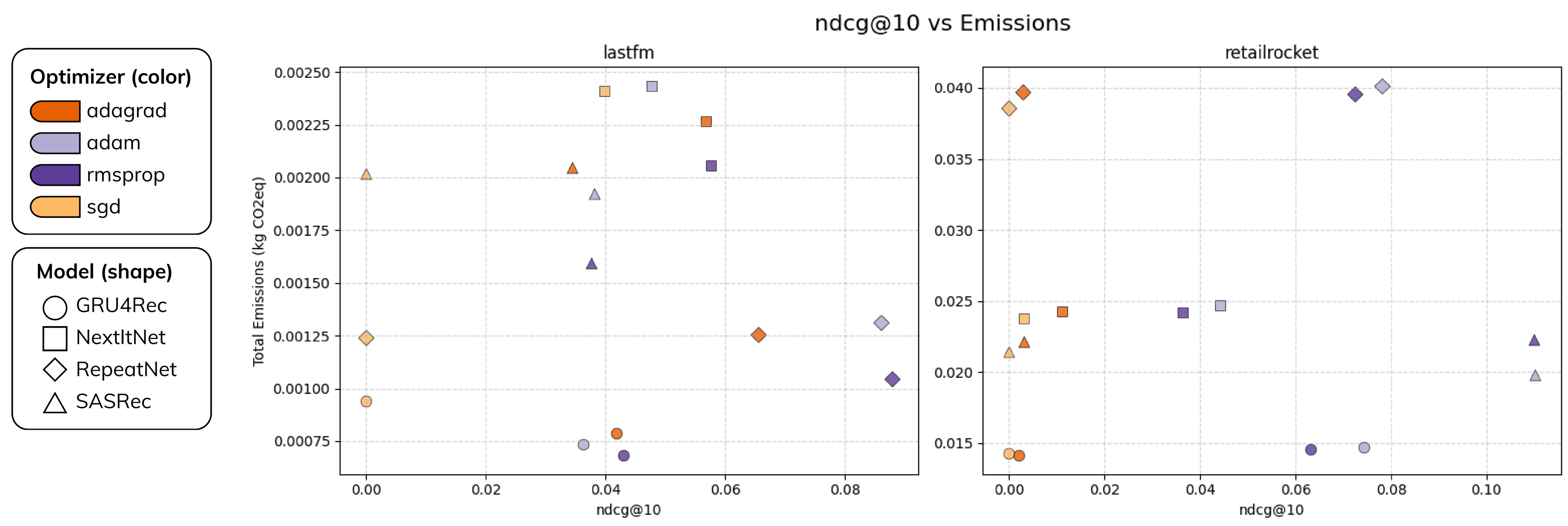
Key Observations

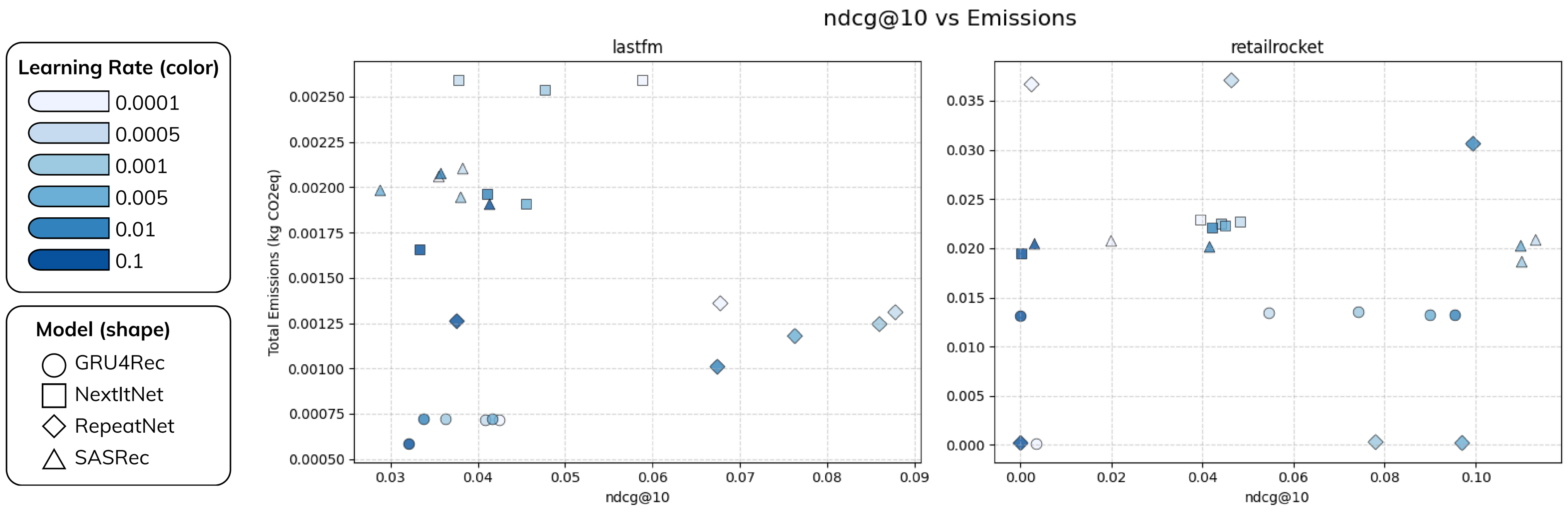
- Training time correlates strongly with total emissions, but not with model performance.
- Caser and NextItNet (convolutional models) show higher emissions and worse performance.
- GRU4Rec achieves low performance but also low emissions.
- SASRec and RepeatNet present the best emissions–performance trade-off despite using different architectures. This aspect warrants further investigation.
- RepeatNet performs better on datasets with frequent item repetitions, while SASRec excels in datasets without strong repetition patterns.

Hyperparameter Analysis

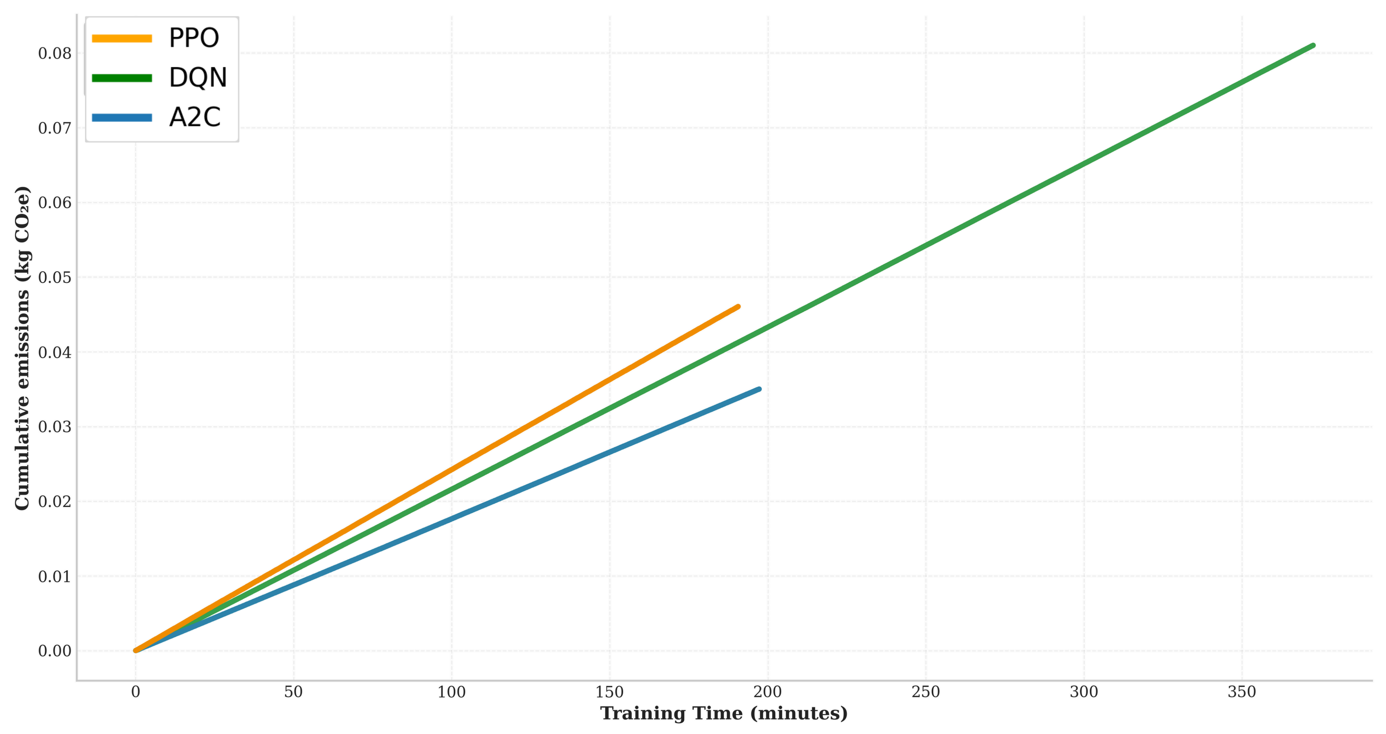
We performed a sensitivity analysis on optimizer choice and learning rate on each model and observed clear differences in both emissions and performance as these parameters were ablated.

- Across optimizers, Adam and RMSProp consistently provided the best emission–performance balance.
- In the learning rate experiments (run with Adam), mid–low values achieved the strongest results.
- Complementary experiments revealed that Adagrad and SGD often perform better with higher learning rates, emphasizing that optimizer and learning-rate tuning must be considered jointly.
- The **choice of model architecture** still has a larger impact on total emissions than hyperparameters alone, reinforcing the importance of selecting efficient models and implementations.





Reinforcement Learning Agents on Atari



We analyze the reinforcement learning agents DQN, PPO, and A2C across 11 different Atari environments to evaluate both their predictive performance and energy consumption. **Clear differences in emission rates emerge, highlighting the importance of considering specific computational characteristics when optimizing for sustainability.**

While DQN typically generates higher total emissions in most environments due to its longer training time to complete the timesteps, PPO exhibits the highest emission rate per unit of time, whereas A2C shows the lowest.

These results underscore the **need to evaluate both total runtime and emission rate**, since an algorithm like PPO would emit even more than DQN if both required the same training time.

Conclusions and Future Steps

Our preliminary results suggest that the **carbon footprint of recommender systems can be substantially reduced through simple changes in architecture and hyperparameters.** Ongoing work includes exploring model characteristics such as hidden size, loss function, and number of layers, as well as integrating reinforcement learning heads into sequential recommenders to evaluate how reinforcement components affect the emissions–performance balance. These insights may support future guidelines for sustainable recommendation research.

Supplementary Material

Contact
Alejandro Plaza: aplaza2@uc.cl
Denis Parra: dparras@uc.cl

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

CENIA — CENTRO NACIONAL DE INTELIGENCIA ARTIFICIAL

HAI —VIS