Netflix Summer 2024 Machine Learning Internships - ML Areas

| Large, Multi-Modal, Multi-Task Models

Large, multi-modal, and multi-task neural network models have shown their value in various applications. We will continue exploring those in the context of Netflix's business as well. Some examples of internship topics are model quantization, distillation, pruning, approximations, etc., Combining media understanding with user behavior understanding, multimodal content understanding, learning tokenized entity representations, scaling training, and serving of large models (especially transformers), multi-task and multi-level learning, etc.

| Generative Language Models in Recommendations

Advances in large language models have opened up new capabilities for building better recommendation experiences. Some examples of internship topics are personalized recommendations explanations, conversational recommendations, personalized LLMs, grounding LLMs in Netflix's domain-specific facts, prompt engineering techniques, etc.

| Reinforcement Learning and Bandits

Because recommender systems can only offer up a limited set of suggestions in each session and receive very sparse feedback, it is natural to look at them through the lens of bandits and reinforcement learning. Making these approaches work well within the personalization domain, however, has unique challenges due to the very large and evolving action spaces, high-dimensional state spaces, and the need for off-policy training and evaluation. Thus, we are conducting research in this area to push forward what is possible and develop better algorithms for long-term optimization.

Examples of internship topics are RL for messaging, time and sequence RL, RLHF-inspired approaches, inverse RL to learn reward function, reward shaping, systematic exploration of member preferences, better system explore strategies and handling of new items, theoretically understanding challenges in adopting RL for recommendations, leveraging additional signals for RL, reward-robust RL, dealing with exploration contamination, etc.

| Evaluation and Causal ML

Some example internship topics in this area are online/offline metrics (and rewards) alignment, causal methods, automated detection of patterns of poor performance, better OPE, metrics for interactive systems (adaptive), metrics for explanations, more robust recommendation quality metrics, simulation system for user behavior, better online experimentation strategies, qualitative feedback, etc.

| General Machine Learning Applications

This includes converting business problems to machine learning problems and building machine learning applications for solving those. This would provide you with valuable practical ML experience in an industry setting.

| Machine Learning Infrastructure

The ML Platform at Netflix is responsible for maximizing the impact of all machine learning practitioners at Netflix. The platform comprises systems, solutions, libraries, and services that touch all aspects of the full machine learning lifecycle from data access, feature engineering, model development, tuning and evaluation, and model inference and serving. Working in this area you will get an opportunity to explore key engineering problems that can advance the ML Platform in innovation areas like AutoML or Bandits and Reinforcement Learning infrastructure.

| Multimedia Processing, Understanding, and Generation

The goal of multimedia processing is to provide the best quality of experience for Netflix members. Example topics in this area include image/video compression, super-resolution, developing deep models to predict subjective quality of processed media, and optimizing adaptive streaming experience end-to-end. Understanding and generation includes Computer Vision, NLP, and Audio Understanding. The amazing content library we love on Netflix comes together due to numerous creative & technical decisions made just the right way: deciding what content to commission, how to produce it, and how to promote it are all critical questions that require a deep & thorough understanding of existing and upcoming content. Internship projects could be around using the latest Computer Vision, Audio ML, Graphics, Image/Video Processing, Multimodal ML, and Generative AI techniques to help characterize our content, help artists during the production process, as well as to help synthesize creative artifacts that connect our titles to our members.