Building Intelligent Recommender Systems

Deep learning-based recommender systems are the secret ingredient behind personalized online experiences and powerful decision support tools in retail, entertainment, healthcare, finance, and other industries.

Recommender systems work by understanding the preferences, previous decisions, and other characteristics of many people. For example, recommenders can predict the types of movies an individual will enjoy based on the movies they've previously watched and the languages they understand. Training a neural network to generalize this mountain of data and quickly provide specific recommendations for similar individuals or situations requires massive amounts of computation, which can be accelerated dramatically by GPUs. Organizations seeking to provide more delightful user experiences, deeper engagement with their customers, and better informed decisions can realize tremendous value by applying properly designed and trained recommender systems.

This workshop covers the fundamental tools and techniques for building highly effective recommender systems, as well as how to deploy GPU-accelerated solutions for real-time recommendations.

Learning Objectives

By participating in this is workshop, you'll learn how to:

- Build a content-based recommender system using the open-source cuDF library and Apache Arrow
- Construct a collaborative filtering recommender system using alternating least squares (ALS) and CuPy
- Design a wide and deep neural network using TensorFlow 2 to create a hybrid recommender system
- Optimize performance for both training and inference using large, sparse datasets
- Deploy a recommender model as a high-performance web service

Workshop Information and Prerequisites:

Duration:	8 hours
Price:	**Contact us for pricing.**
Prerequisites:	- Intermediate knowledge of Python, including understanding of list comprehension - Data science experience using Python - Familiarity with NumPy and matrix mathematics Suggested resources to satisfy prerequisites: **Python Tutorial**, **Kaggle Learn Machine Learning Course**, **NumPy Tutorial**
Tools, libraries, and frameworks:	**CuDF**, **CuPy**, **TensorFlow 2**, and **NVIDIA Triton™ Inference Server**
Assessment type:	Skills-based coding assessments evaluate students' ability to debug and correct production-quality recommendation pipelines.
Certificate:	Upon successful completion of the assessments, participants will receive an NVIDIA Deep Learning Institute certificate to recognize their subject matter competency and support professional career growth.
Hardware/software requirements:	Desktop or laptop computer capable of running the latest version of Chrome or Firefox. Each participant will be provided with dedicated access to a fully configured, GPU-accelerated workstation in the cloud.
Language:	English

Sample Workshop Outline

Introduction (15 mins)	> Meet the instructor. > Create an account at **courses.nvidia.com/join**
Matrix-Based Recommender Systems (120 mins)	**Implement collaborative filtering with singular value decomposition (SVD):** > Read sparse data into a GPU using **CuPy**. > Perform ALS efficiently with NumPy broadcasting rules. > Build a content-based filter with **cuDF**.
Break (60 mins)	
Training Wide and Deep Recommenders (120 mins)	**Build a wide and deep network using TensorFlow 2:** > Build a deep network using Keras. > Build a wide and deep network using TensorFlow feature columns. > Efficiently ingest training data with tf.data. > Case study 1: See real-world examples of recommender system model architectures.
Break (15 mins)	
Challenges of Deploying Recommendation Systems to Production (120 mins)	**Deploy a recommender system in a production environment:** > Acquire a trained model configuration for deployment. > Build a container for deployment. > Deploy the trained model using NVIDIA Triton Inference Server.
Final Review (15 mins)	> Review key learnings and answer questions. > Learn to build your own training environment from the DLI base environment container. > Complete the assessment and earn a certificate. > Take the workshop survey. > Case study 2: Review real-world challenges of at-scale recommender systems.

Why Choose NVIDIA Deep Learning Institute for Hands-On Training?

- Access workshops from anywhere with just your desktop/laptop and an internet connection. Each participant will have access to a fully configured, GPU-accelerated workstation in the cloud.
- Obtain hands-on experience with the most widely used, industry-standard software, tools, and frameworks.
- Learn to build deep learning and accelerated computing applications for industries, such as healthcare, robotics, manufacturing, accelerated computing, and more.
- Gain real-world experience through content designed in collaboration with industry leaders, such as the Children's Hospital of Los Angeles, Mayo Clinic, and PwC.
- Earn an NVIDIA Deep Learning Institute certificate to demonstrate your subject matter competency and support your career growth.

For the latest DLI workshops and trainings, visit **www.nvidia.com/dli**

For questions, contact us at **nvdli@nvidia.com**