Streamline AI/ML Development for Data Scientists with Workstations Powered by Intel

What do data scientists need?

Data scientists occupy a unique role within development teams building AI solutions. At the start of any AI project, the data preparation, model evaluation, preprocessing, and data exploration are the core tasks.

Developing the AI model is a patient, iterative, thoughtful process—including testing different strategies and models, and predicting performance.

Data scientists need frameworks and applications that excel at data exploration; extract, transform, and load (ETL) operations; and visualization tasks, favoring single-node deployments.

Interacting with massive datasets and maintaining a single-node memory span rule out cloud deployments. Workstations enabled for data locality minimize latency issues.

builders.intel.com/ai

Build the ultimate workstation geared to data science efforts

Data scientists require highly interactive systems that can handle massive volumes of data, using tools designed for single-node processing. The Initial Phase in the workflow (see Figure 1) occupies about 80% of the overall effort and benefits from a workstation optimized for key tasks. The Intel portfolio offers hardware, software, libraries, accelerators, frameworks, and toolkits with the ideal capabilities for data science projects.

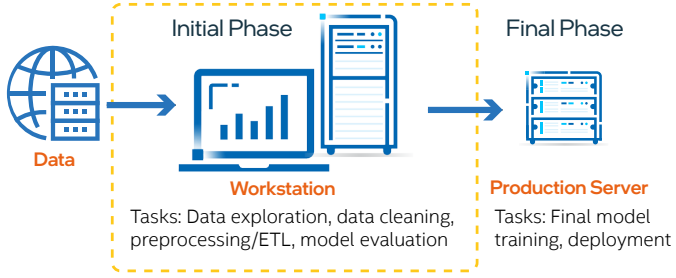


Figure 1. Typical AI/ML workflow.

Choose purpose-built hardware for specialized data science tasks

A slate of processors infused with AI capabilities power the workstations offered by Intel, coupled with persistent memory options and fast SSDs. Compared to high-performance computing cluster nodes or premium laptops, workstation memory and drives can be effectively configured for larger volumes of data.

Capitalize on open standards and a cross-architecture development platform

OEMs and system integrators providing workstations configured for data scientist requirements can take advantage of optimized open standards software tuned for AI acceleration. System building is streamlined by toolkits that support a unified programming model to deliver powerful, cross-platform interoperability.

Take advantage of extensive ecosystem resources

Minimize the cost and complexities of designing AI solutions for data scientists by tapping into a vast array of resources provided by the dynamic ecosystem of the Intel® AI Builders program. Members gain access to tools and technologies to accelerate AI adoption and opportunities for co-marketing and matchmaking with technology leaders. Learn more at builders.intel.com/ai.

Earn the trust and confidence of data scientists by offering workstation solutions keyed to their requirements.

Guidelines for building those workstations appear on page two.

Successful AI solutions are grounded in hardware. Intel® Xeon® processors and Intel Core™ processors power workstations with deep capabilities for diverse uses. No GPU required.

OEMs and systems integrators selling to the data science sector have a unique opportunity to offer purpose-built workstations that are cost effective and scale to ingest massive datasets.

Entry level

Mobile workstations

Configuration for data science enthusiasts.
Single-socket Intel Core i9 -10900k processor, 3.7 GHz, 10 cores/20 threads

- Cache sizes: 20 MB, 2.5 MB, 320 KB
- 128 GB memory (4 x 32 GB)
- 2TB SSD

Primary use cases

Best for basic data science projects, meeting baseline memory and storage requirements for budget-sensitive uses. Intel Core processors excel where workloads scale with raw clock speed (not requiring AVX-512 speeds).

Approximate price as configured: USD 6,000[1]

Mid-tier

Mid-tier workstations

Balanced core count and frequency.
Single-socket Intel Xeon W 2295 processor, 3.0 GHz, 18 cores/36 threads

- Cache size: 24.75 MB, 18 MB, 1 MB
- 4 memory channels
- 512 GB memory (8 x 64 GB)
- 2 TB SSD
- Intel Optane™ SSD 905P Series (960 GB) (AIC PCIe x 4 3D XPoint)

Primary use cases

Achieves an equitable balance between core count and processor frequency, providing cost-effective performance within moderate heat and power parameters.

Approximate price as configured: USD 16,000[1]

Top-tier

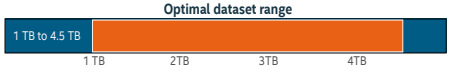
Top-tier workstations

Demanding applications with broad memory span.
Dual-socket Intel Xeon Gold 6258R processor, 2.7 GHz, 28 cores/56 threads

- Cache size: 38.5 MB, 28 MB, 1.75 MB (Intel Xeon processor L2/L3 cache hierarchy)
- 1024 GB (1 TB) memory (16 x 64 GB DDR4 ECC RDIMM)
- 2 TB SSD

Primary use cases

Suits applications in which requiring the memory span exceeds 3 TB, ranging as high as 6 TB. Also favors applications in which server-based optimization and core scaling are essential. This configuration requires balancing benefits against the available power budget.



Approximate price as configured: USD 30,000[1]

The workstation configurations and use cases shown are a representative sampling.

Explore other choices at Workstations Powered by Intel.

Preload optimized libraries and frameworks from Intel to help data scientists improve their workflow.

Intel oneAPI Toolkits

Unify coding efforts with a single programming model that delivers exceptional cross-architecture performance

Built on open standards, **Intel oneAPI** simplifies development and deployment of data-centric workloads. The product family includes compilers, performance libraries, analyzers, debuggers, and more. Unifies coding for CPUs, GPUs, FPGAs, and other accelerators. Some of the components that are particularly valuable to data scientists are described in more detail in the following subsections.
Check out the **Installation Guide for Intel oneAPI Toolkits** for Linux, Windows, and macOS.

Intel Distribution for Python

Accelerate compute-intensive applications—from data analytics to machine learning—that use NumPy, SciPy, scikit-learn, and more

An essential tool for data scientists, this distribution delivers faster application performance on Intel platforms. Implement and scale production-ready algorithms for scientific computing and machine-learning workloads, and optimize performative with native Intel Performance Libraries. **Download the free package** and unleash a faster Python on your data.

Intel AI Analytics Toolkit

Boost performance across data science and AI pipelines with this toolkit, built using oneAPI libraries

Enable significant acceleration for data analytics and machine-learning workflows, as well as direct access to Intel analytics and AI optimizations across the entire AI pipeline, from preprocessing through machine learning. **Download the Intel oneAPI AI analytics Toolkit**.

Intel oneAPI HPC Toolkit

Enhance high-performance computing applications with data-centric libraries, a powerful compiler, and advanced analysis tools

This add-on to the Intel oneAPI Base Toolkit provides valuable tools for data scientists building, analyzing, optimizing, and scaling HPC applications. The toolkit incorporates the latest techniques in vectorization, multithreading, multi-node parallelization, and memory optimization. **Download the Intel oneAPI HPC Toolkit**.

Intel's commitment to enhancing the AI journey for developers extends from the cloud to the edge and to end devices.

To discover ways to improve the AI journey for customers, visit builders.intel.com/ai.

[1]System pricing assumes a baseline configuration for BOXX Technologies model GoBOXX 17 FHD MXL (mobile workstation), APEXX W3/W4L (mid-tier workstation), and APEXX D4 (top-tier workstation). Visit www.boxx.com for more information.