ActiveState®

ACTIVEPYTHON
FOR MACHINE LEARNING

FASTER ALGORITHMS
Pre-integrated with Intel MKL and pre-optimized for speed

LESS DATA WORK
Automate the bookkeeping of metadata with Pandas

BETTER COLLABORATION
Share code & annotations via Jupyter

DRIVE RESULTS
Rapidly build & evaluate models with Keras

DECREASE TIME TO MARKET
Get precompiled Python & machine learning packages

REDUCE RISK
Python security scanned & backed by SLA-based support

ENSURE COMPLIANCE
Vetted pre-bundled 3rd-party packages

NO VENDOR LOCK-IN
100% compatible with community open source Python

Transform Data into Knowledge Using Machine Learning

Machine Learning is fast becoming a key, strategic initiative within the enterprise, allowing you to derive insight from all the data you've been collecting, and then leverage it to create differentiation in the marketplace.

Machine learning can help you solve business problems utilizing mathematical models that extract knowledge from data. Falling under the discipline of Data Science, machine learning can aid enterprises in achieving everything from sales & marketing goals to strategic & financial planning to risk & fraud detection. Machine learning models separate signal from noise in real time to help you interactively spot trends and anomalies in large data sets better and faster than simple statistical analyses.
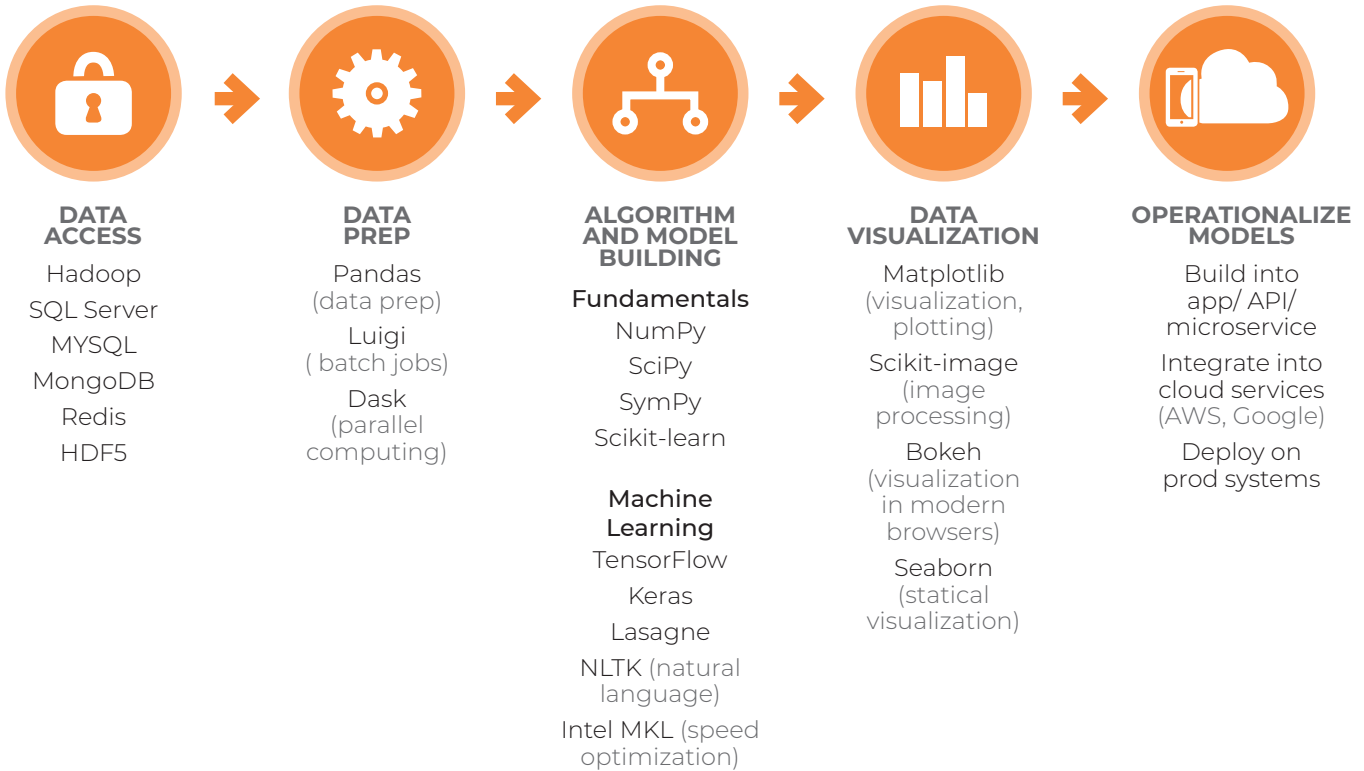
GETTING STARTED IN ML
While there are numerous vendors of proprietary data science, big data, and scientific computing applications, building a custom machine learning solution can provide you with greater agility and control, especially if you have unique data that grants you key differentiation in the marketplace. However, operationalizing machine learning models by embedding them either in a production application or business process remains a major challenge.

Python has emerged as a key productizing tool for machine learning projects due to its high productivity (being a scripting language), and wealth of third-party data science components, but it can be orders of magnitude slower than native code. Machine learning models are derived from numerical algorithms that are typically computationally expensive, which makes performance a key consideration. More importantly, for realistic scientific modeling it's critical to scale from a prototyping environment, which is typically a laptop or a workstation, to multiple nodes.

ACTIVEPYTHON
FOR MACHINE LEARNING

DATA ACCESS	DATA PREP	ALGORITHM AND MODEL BUILDING	DATA VISUALIZATION	OPERATIONALIZE MODELS
Hadoop	Pandas (data prep)	**Fundamentals**	Matplotlib (visualization, plotting)	Build into app/ API/ microservice
SQL Server	Luigi (batch jobs)	NumPy	Scikit-image (image processing)	Integrate into cloud services (AWS, Google)
MYSQL	Dask (parallel computing)	SciPy	Bokeh (visualization in modern browsers)	Deploy on prod systems
MongoDB		SymPy	Seaborn (statical visualization)	
Redis		Scikit-learn		
HDF5		**Machine Learning**		
		TensorFlow		
		Keras		
		Lasagne		
		NLTK (natural language)		
		Intel MKL (speed optimization)		

Machine learning projects start with the data. Typically, 80% of the work is cleaning up the data, feeding it to your algorithms and training the machine learning component. If you've done a good job normalizing the data, you'll get convergence and a model you can use.

ActivePython includes open source community packages like Pandas to help with the data pre-processing. Packages like TensorFlow, and Keras, as well as scikit-learn, provide the algorithms, additional libraries, computational power and user-friendly control to develop the learning stages. A key bottleneck in any machine learning project is the processing of algorithms. ActivePython incorporates Intel's Math Kernel Library (MKL), which takes advantage of multiple cores and vector registers to accelerate basic linear algebra operations and solvers, Fast Fourier Transforms (FFTs), arithmetic and transcendental operations, and more. This means mathematical routines and model training run faster, so you can get your project to market faster.

With a single, consistent Python environment across your data science and programming teams, productizing your machine learning model is less like throwing it over the wall to coders and more like passing the baton. Python shines as a rapid prototyping environment, providing web frameworks like Django and Flask that allow your developers to incorporate your learning models in web apps, APIs, etc., and then scale them out in production using integrated cloud tools for AWS and Google.

While open source Python provides many of the tools and libraries for Machine Learning, high-value staff can end up wasting days on the low-value work of installing and configuring packages before they are able to start writing algorithms. ActivePython not only comes pre-compiled with the most popular open source packages for machine learning, but is also pre-optimized for compatibility and speed, ensuring data scientists and application development teams can be productive right out of the box.