



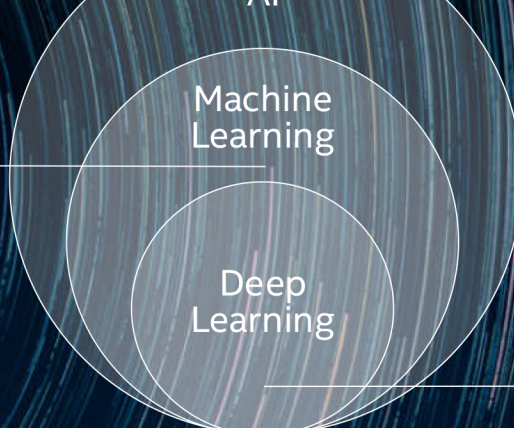
# AI 101:

## A QUICK GUIDE TO CHOOSING THE RIGHT AI APPROACH FOR YOUR BUSINESS NEEDS

As **artificial intelligence (AI)** techniques have rapidly developed, many enterprises look to apply it to a diversity of problems. However, the term “artificial intelligence” can be interpreted in different ways across a wide range of approaches. **This quick guide is intended to help you determine where and how you can incorporate AI to address your business needs.**

### WHAT IS AI?

- Regression
- Classification
- Clustering



- Image processing
- Speech processing
- Natural language processing
- Recommender systems
- Adversarial networks

AI refers to a broad class of systems that enable machines to mimic human cognitive abilities. Machine learning (ML) is the most common way to achieve artificial intelligence today, and deep learning (DL) is a subset of machine learning.

### ASSESS YOUR BUSINESS REQUIREMENTS

Begin by defining your business use case – what workflow, prediction, or process do you need to automate or gain insights from? What are the characteristics of the data you'll be using in terms of format, volume, when and how often it needs to be analyzed? This will help determine the most appropriate type of AI algorithm for your business.

Business Use Case	Data Characteristics	Type of AI	Class of AI Algorithm
Discover the relationship between variables	Structured, batch analysis	Machine learning	Regression
Predict categorical class labels or classify data	Structured, batch analysis	Machine learning	Classification
Uncover meaningful relationships between variables	Structured, batch analysis	Machine learning	Clustering
Classify images, locate and classify object(s) in an image	Unstructured visual data (video, MRI scans etc.), batch or real-time analysis, small datasets	Deep learning: Image recognition, object detection	Convolutional neural network (CNNs)
Extract content and meaning from text	Unstructured text (email, Office documents, social media posts etc.), unstructured audio/video content, batch or real-time analysis	Deep learning: Natural language processing (NLP), language translation, speech recognition	Long short-term memory (LSTM), recursive neural network (RNN)
Convert text to speech	Unstructured text, real-time analysis	Deep learning: Text-to-speech	Generative adversarial network (GAN)
Recommend ads, search, apps etc.	Structured customer data, unstructured text, real-time analysis	Deep learning: Recommendation engine	Multilayer perception (MLP)
Create images like training data	Unstructured image and video, real-time analysis	Deep learning: Data generation	GAN with CNN
Learning from feedback on action	Structured and unstructured, batch or real-time analysis	Deep learning: Reinforcement learning	CNN or CNN+RNN

### GET STARTED USING FAMILIAR TECHNOLOGY

Once you've identified your first project, it's usually possible to start on your AI journey right away, using your existing IT infrastructure in many cases. Any AI solution needs to take into account three key areas:



Hardware



Software



Ecosystem

Intel® technology is designed to help you make the most of your entire AI solution stack:

Hardware		Compute: 2nd generation Intel® Xeon® Scalable servers with Intel® Deep Learning Boost (Intel® DL Boost) is the foundation for AI and offers <b>up to 30x performance improvement</b> <sup>1</sup> for AI workloads. Intel also has a flexible edge-to-cloud portfolio to meet your growing AI needs
Software		Optimized AI Software Tools: Common ML and DL frameworks, toolkits and libraries are optimized to accelerate AI on Intel® architecture. For example Intel and Google's TensorFlow optimization delivers <b>up to 3.75x inferencing speed up</b> when using Intel DL Boost <sup>2</sup>
Ecosystem		Simplified AI deployment: Implement and scale AI solutions in a way that works for your organization, with ready-made solutions from Intel AI Builders members; or build your own with ecosystem support and pre-verified Intel Select Solutions for common AI workloads

### LEARN MORE

Still unsure where to start? Explore some real-world examples of AI in action to see what's possible:

- Learn how to [accelerate your AI adoption](#) with minimal cost and complexity
- See how JLK Inspection is helping medical professionals [accelerate screening](#) and enhance patient care with the Intel® Distribution of OpenVINO™ toolkit
- Follow the journey of Kongsberg Maritime as it [shifted its AI solutions from GPU- to CPU-based infrastructure](#), and increased its performance
- Discover how Intel-optimized AI frameworks are helping [convert raw data into market insights](#) to guide decision making and inform investors in speciality industries
- Explore how conversational AI using Intel® technologies are helping contact center agents [enhance the customer experience](#)

For more information on Intel Enterprise Solutions for AI, please contact us today.

Enterprise Technology International



<sup>1</sup> 30x inference throughput improvement on Intel® Xeon® Platinum 9282 processor with Intel® DL Boost : Tested by Intel as of 2/26/2019. Platform: Dragon rock 2 socket Intel® Xeon® Platinum 9282/56 cores per socket, HT ON, turbo ON, Total Memory 768 GB (24 slots/ 32 GB/ 2933 MHz), BIOS:SE5C620.86B.0D.01.0241.112020180249, Centos 7 Kernel 3.10.0-957.5.1.el7.x86\_64, Deep Learning Framework: Intel® Optimization for Caffe version: HYPERLINK "https://github.com/intel/caffe%20d554cbf1" https://github.com/intel/caffe d554cbf1, ICC 2019.2.187, MKL DNN version: v0.17 (commit hash: 830a10059a018cd2634d94195140cf2d8790a75a), model: https://github.com/intel/caffe/blob/master/models/intel\_optimized\_models/int8/resnet50\_int8\_full\_conv.prototxt, BS=64, No datalayer governor set to "performance" via intel\_pstate driver, 384GB DDR4-2666 ECC RAM, CentOS Linux release 7.3.1611 (Core), Linux kernel 3.10.0-514.10.2.el7.x86\_64, SSD: Intel® SSD DC 53700 Series (800GB, 2.5in SATA 6Gb/s, 25nm, MLC).Performance measured with: Environment variables: KMP\_AFFINITY=granularity=fine,compact, OMP\_NUM\_THREADS=56, CPU Freq set with cpupower frequency-set -d 2.5G -u 3.8G -g performance. Caffe: (http://github.com/intel/caffe/), revision f96b759f71b2281835f690af267158b82b150b5c. Inference measured with "caffe time --forward\_only" command, training measured with "caffe time" command. For "ConvNet" topologies, synthetic dataset was used. For other topologies, data was stored on local storage and cached in memory before training. Topology specs from https://github.com/intel/caffe/tree/master/models/intel\_optimized\_models (ResNet-50), Intel C++ compiler ver. 17.0.2 20170213, Intel MKL small libraries version 2018.0.20170425. Caffe run with "numactl -l".

<sup>2</sup> 3.75x improvement with AI Inferencing Intel Select Solution. The solution was tested with KPI Targets: OpenVINO/ ResNet50 on INT8 on 02-26-2019 with the following hardware and software configuration: Base configuration: 1 Node, 2x Intel® Xeon® Gold 6248; 1x Intel® Server Board S2600WFT; Total Memory 192 GB, 12 slots/16 GB/2666 MT/s DDR4 RDIMM; HyperThreading: Enable; Turbo: Enable; Storage(boot): Intel® SSD DC P4101; Storage(capacity): At least 2 TB Intel® SSD DC P4610 PCIe NVMe; OS/Software: CentOS Linux release 7.6.1810 (Core) with Kernel 3.10.0-957.el7.x86\_64; Framework version: OpenVINO 2018 R5 445; Datasets/sample image from benchmark tool; Model topology: ResNet 50 v1; Batch Size: 4; nireq: 20. The solution was tested with KPI Targets: TensorFlow/ ResNet50 on INT8 on 03-07-2019 with the following hardware and software configuration: Base configuration: 1 Node, 2x Intel® Xeon® Gold 6248; 1x Intel® Server Board S2600WFT; Total Memory 192 GB, 12 slots/16 GB/2666 MT/s DDR4 RDIMM; HyperThreading: Enable; Turbo: Enable; Storage(boot): Intel® SSD DC P4101; Storage(capacity): At least 2 TB Intel® SSD DC P4610 PCIe NVMe; OS/Software: CentOS Linux release 7.6.1810 (Core) with Kernel 3.10.0-957.el7.x86\_64; Framework version: intelaipg/intel-optimized-tensorflowPR25765-devel-mkl; Dataset: Synthetic from benchmark tool; Model topology: ResNet 50 v1; Batch Size: 80

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit [www.intel.com/benchmarks](http://www.intel.com/benchmarks).

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

0620/JL/CAT/PDF

343566-001EN