EXECUTIVES' GUIDE TO AI

AI using machine learning (ML) has gained popularity in recent years as it promises to revolutionize business operations. For businesses, artificial intelligence (AI) offers numerous benefits that enable them to remain competitive in an ever-evolving world. AI can automate mundane and repetitive tasks, streamline decision-making processes, reduce costs, increase efficiency and accuracy, and improve customer service. It also enables companies to manage their processes more effectively. These are all competitive advantages made possible with AI.

In the medium term, companies that do not adopt this technology will lose out in terms of innovation and customer experience, making it more difficult for them to remain competitive. This could ultimately lead to lost profits and revenue if customers can expect a tangible advantage from companies that adopt AI.

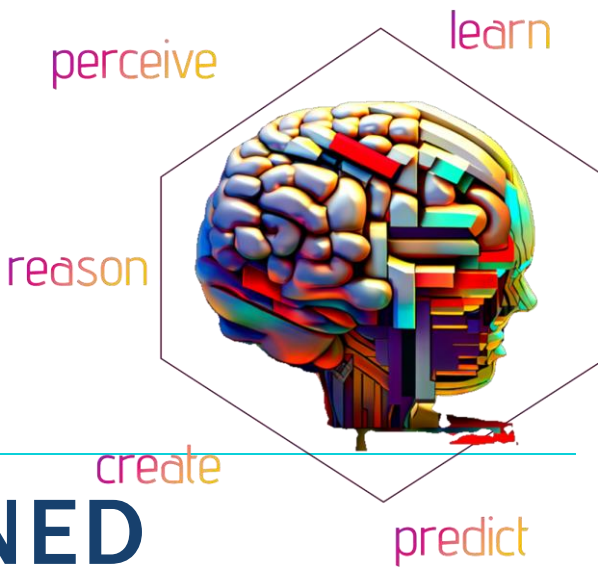
Executives who want to stay ahead of the curve need to adopt new technologies and strategies quickly, but often don't have the time or expertise to process enough information to make the right decisions on behalf of the business.

It is therefore critical that CEOs and CTOs understand both the benefits and challenges of AI so they can make informed decisions about how best to integrate AI into business operations.

This guide explains AI, the different types of AI, and common business problems that can be solved with AI and ML in simple terms.

AI DEFINED

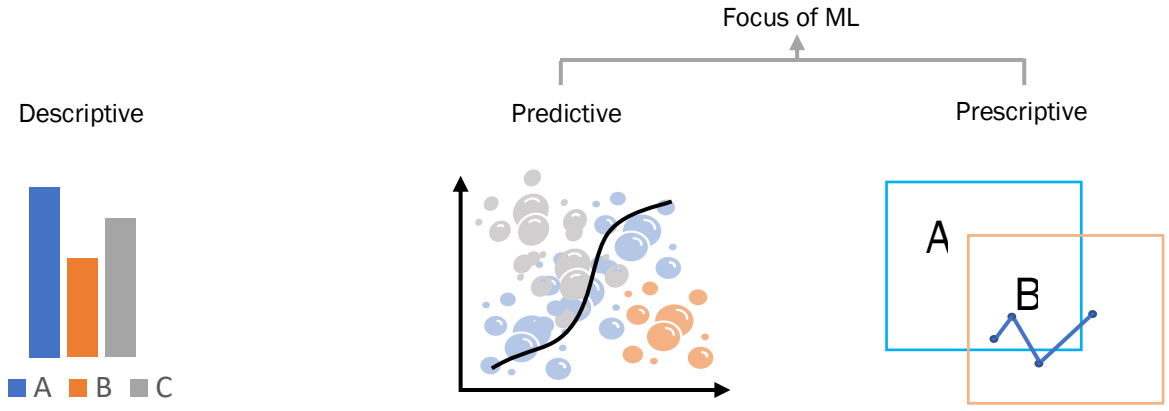
AI denotes a machine's capability to do things we normally think only humans can, such as seeing, thinking, learning and problem-solving. Technologies that enable machines to do these things include robotics and autonomous vehicles, computer vision, language processing, virtual agents and machine learning.



ML DEFINED

The application of machine learning to sizable data sets is behind the most recent advancements in AI. Rather than receiving explicit programming instructions, machine learning systems recognize patterns and learn how to make predictions and suggestions by analyzing data and experiences. Additionally, the algorithms adjust themselves in response to new input data and experiences to boost their effectiveness over time.

Standard analysis versus ML's identification of insights



- Tells WHAT HAPPENED
- Heavily used in all domains

- Tells WHAT WILL HAPPEN (probabilistic in nature)
- Source of insights for top data-driven organizations

- Tells WHAT TO DO
- Used by leading data and internet service providers

Major types of machine learning

	Supervised Learning	Unsupervised Learning	Reinforcement Learning
WHAT?	An algorithm that learns the link between specified inputs and outputs using training data and feedback from humans (e.g., how the inputs "season of year" and "interest rates" predict home prices).	An algorithm that analyzes input data without being provided a specific output variable (e.g., analyzes customer demographic data to identify buying trends).	An algorithm that learns to complete a task by attempting to maximize its rewards for its actions (e.g., by maximizing the points it receives for raising the returns on an investment portfolio).
WHEN?	You know how to categorize the incoming data and what kind of behavior you need to forecast but you need the algorithm to perform the calculation with fresh data.	You want the algorithm to detect patterns and classify the data for you because you do not know how to categorize the data.	You lack sufficient training data, cannot precisely specify the desired end state or must interact with the environment in order to learn about it.
HOW?	• When predicting home prices, every piece of input data is given a name, such as "season of year," "interest rates," etc. and the output variable is defined as housing prices. • The algorithm is trained on the data to determine the relationship between the input variables and the output. • After training, the method is usually applied to fresh data once the algorithm is sufficiently accurate.	• The algorithm is given unlabeled data, such as a collection of data on client website traffic. • It infers a structure from the data. • The algorithm detects clusters of data that exhibit similar behavior (e.g., groups of customers that make similar purchasing decisions).	• The algorithm changes the environment (e.g., by making a trade in a financial portfolio). • If the activity moves the machine one step closer to maximizing the total rewards (e.g., the highest overall return on the portfolio), it is rewarded. • The algorithm corrects itself over time to optimize itself for the best set of actions.

Supervised learning and business use cases

Algorithm	Business use case
Linear regression	
Modeling the historical link between independent input factors and dependent output variables using a highly interpretable standard approach (which can take any number of different forms) to help predict the future values of the output variables.	• Optimizing pricing points and calculating product-price elasticities. • Understanding the factors that influence how much a product sells, such as competitor prices, distribution, advertising, etc.
Logistic regression	
Extensions of linear regression where the output variable is binary (e.g., only black or white) rather than continuous are used for classification problems (e.g., an infinite list of potential colors).	• Determining and classifying a customer's repayment propensity. • Determining whether a skin lesion is benign or malignant based on its features (size, form, color, etc.).
Linear and quadratic discriminant analysis	
Improves a logistic regression to handle nonlinear issues or issues where changes to the value of the input variables do not have a proportional impact on the output variables.	• Determining the likelihood of closing sales leads. • Classifying customers by their likelihood of making a purchase. • Predicting client churn.
Decision Tree	
Highly interpretable classification or regression model that splits data-feature values into branches at decision nodes (e.g., if a feature is a color, each potential color becomes a new branch) until a final decision output is produced.	• Establishing an HR support system for hiring qualified employees. • Identifying the product features that increase their likelihood of being purchased.
Naïve Bayes	
A classification technique that enables calculating the probability of an event based on knowledge of factors that could influence that event (e.g., an email containing the word "money," has a high likelihood of being spam).	• Determining market perception of a product using sentiment analysis. • Detecting spam emails by establishing classifiers.

Algorithm	Business use case
Support vector machines	
Frequently applied to classification and regression, this method makes the widest possible distinction between classes. It can also be easily generalized to address nonlinear issues.	• Estimating the number of patients for whom a hospital will need to care for a specific period. • Calculating the likelihood that someone will click on an internet advertisement.
Random Forest	
Multiple decision trees are created and a majority vote is taken to determine the output, which is a continuous variable (such as age) for a regression problem and a discrete variable (such as black, white or red) for classification, to increase the accuracy of a simple decision tree.	• Forecasting call volume to make staffing decisions in call centers. • Predicting electricity consumption in a distribution grid.
AdaBoost	
A classification or regression strategy that evaluates several models by how well they predict the outcome while using a large number of models to reach a conclusion.	• Spotting credit card fraud. • Classifying images easily and affordably (e.g., identifying land use from satellite photos for climate change models).
Gradient-boosting trees	
A classification or regression strategy that sequentially produces decision trees, each tree being intended to fix the flaws introduced in the model by the decision tree before it. The predictions from all trees are combined to create the final output.	• Predicting product demand and inventory levels. • Estimating car prices based on attributes (e.g., age, make, model, condition and mileage).
Simple neural network	
A model that can be used to categorize data or determine the relationship between variables in regression issues is made up of three layers: an input layer, a hidden layer in which calculations are performed and an output layer.	• Predicting whether registered users will be willing to pay a specific price for a product. • Estimating the likelihood that a patient will enroll in a healthcare program.

Unsupervised learning and business use cases

Algorithm	Business use case
K-means clustering	
Organizes data into a certain number of groups (k), each of which contains data with a certain set of comparable	• Dividing customers into groups based on distinct characteristics (such as age group) in order to

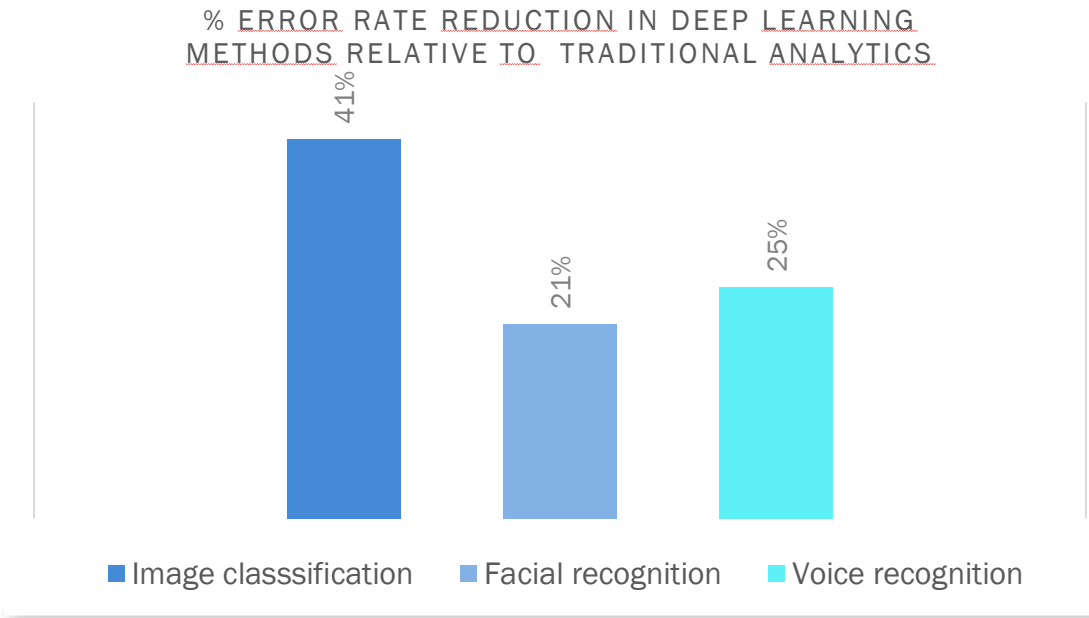
properties (decided by the model rather than by humans beforehand).	target marketing campaigns better or prevent churn.
Gaussian mixture model	
K-means clustering is generalized to allow for larger group size and shape flexibility (clusters).	• Dividing customers into groups based on less distinct characteristics (such as product preferences) in order to target marketing campaigns better. • Dividing employees into groups based on the likelihood of attrition.
Recommender system	
Frequently makes use of cluster behavior prediction to find the crucial information needed for a recommendation.	• Suggesting movies based on the tastes of other customers who share similar characteristics. • Providing news stories that might interest readers based on the article they are currently reading.
Hierarchical clustering	
Creates a classification system by dividing or combining groups along a hierarchical tree.	• Grouping loyalty-card holders into more micro segmented groups. • Grouping customer mentions of terms in social media data to facilitate product use and development.

Reinforcement learning and business use cases

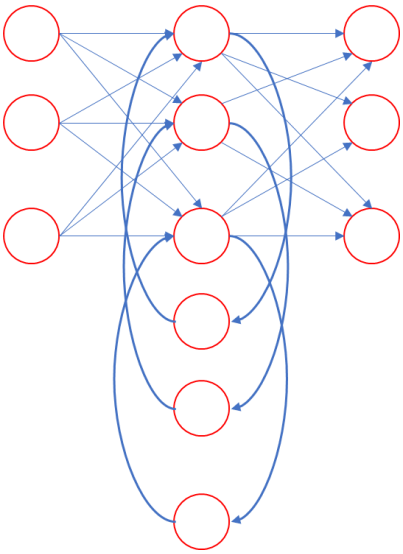
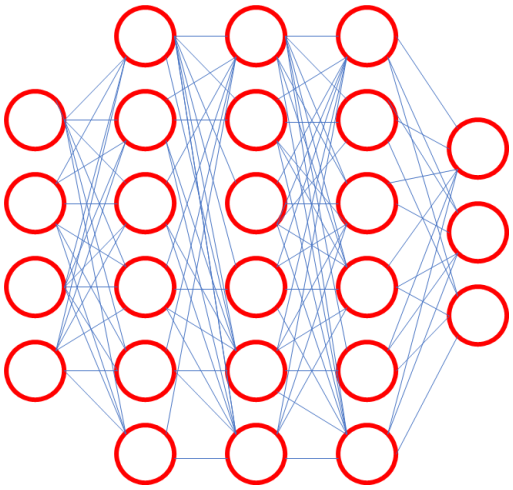
- Improving options portfolio management.
- Adapting energy grid loads to changing demand cycles.
- Using robots to pick and stock inventory.
- Optimizing the performance or functionality of self-driving cars.
- Optimizing real-time pricing for online auctions of limited-supply products.

DEEP LEARNING

Deep learning processes a wider variety of data sources. It requires less data pre-processing by humans and often provides more accurate predictions than traditional machine learning. Neuronal networks with which Deep Learning is possible consist of interconnected layers of computer-built cells called neurons. A network can analyze enormous amounts of input data in numerous layers, each of which learns increasingly complex features of the data. The network can then decide what to do with the data once it has learned whether the data is correct or not, and it can apply what it has learned to decide what to do with new data. For example, if it is familiar with the appearance of an object, it can identify that object in new images.



Key deep learning models and business use cases

	Recurrent neural network	Convolutional neural network
		
WHAT?	It is a multilayer neural network that can learn data sequences and output a number or another sequence by storing information in context nodes.	It is a multilayer neural network with a unique architecture made to extract more complicated elements from the data at each layer to determine the output.
WHEN?	When you are using time-series information or sequences (e.g., audio recordings or text).	When you need to infer information from an unstructured data source (such as photos).
HOW?	In the example, filling in the blank space in the question, "Are you free _____?" 1. A command indicating the beginning of a sentence is given to a neuron in a recurrent neural network (RNN). 2. After receiving the word "are", the neuron produces a vector of numbers that feeds back into itself to help the neuron "remember" that it (and it alone) just received the word "are". The same happens when it receives "you" and "free", the neuron's state being updated after each word. 3. After receiving the word "free", the neuron estimates the likelihood of each	In image processing, for instance, 1. When given an image, such as the letter "A," the convolutional neural network (CNN) interprets it as a group of pixels. 2. It locates distinctive elements in the model's inner hidden layers, such as the several lines comprising the letter "A." 3. The CNN can now identify the distinctive elements that comprise the letter "A" in a new image if it finds them there.

word in the English language being able to complete the sentence. A properly trained RNN will output and select "tomorrow", one of the highly probable words, to complete the sentence.

- Creating analyst reports for securities traders.
- Providing language assistance.
- Monitoring visual changes of a site following a disaster to evaluate potential damage claims (in conjunction with CNNs).
- Determining the likelihood that a credit card transaction is fraudulent.
- Creating captions for pictures.
- Setting up intelligent chatbots that can handle more complex consumer requests and questions.
- Using medical imaging to identify medical conditions.
- Identifying companies' logos in social media activity to explore strategic partnerships, brand unification and co-branding initiatives
- Using visuals to analyze consumer brand perception
- Using photos to spot defective items on a production line.