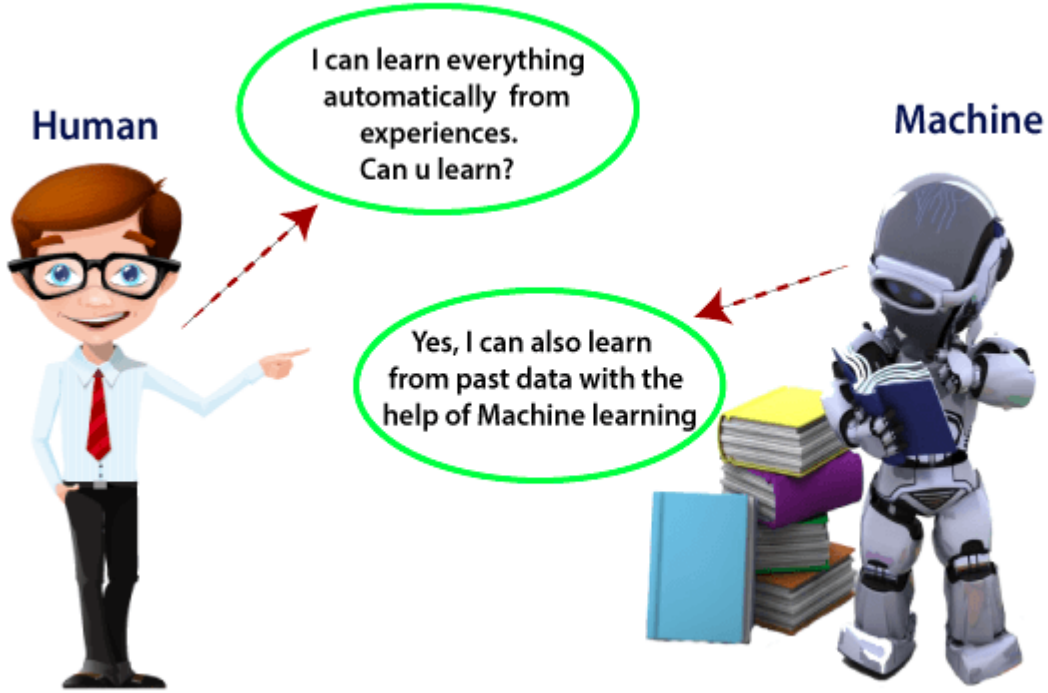
Introduction to Machine Learning

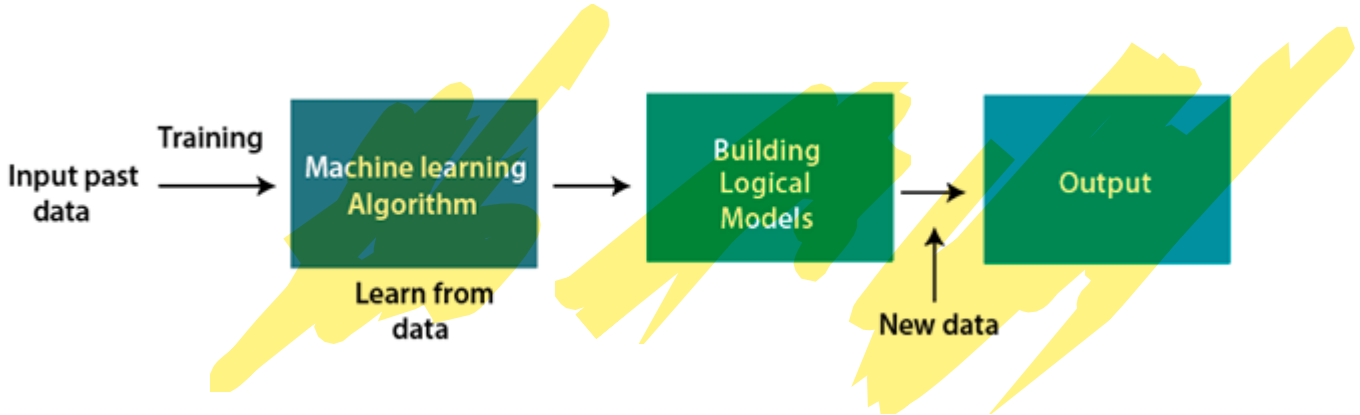
Machine Learning

- In the real world, we are surrounded by humans who can learn everything from their experiences with their learning capability, and we have computers or machines which work on our instructions.
- But can a machine also learn from experiences or past data like a human does? So here comes the role of Machine Learning.
- Machine Learning is said as a subset of artificial intelligence that is mainly concerned with the development of algorithms which allow a computer to learn from the data and past experiences on their own



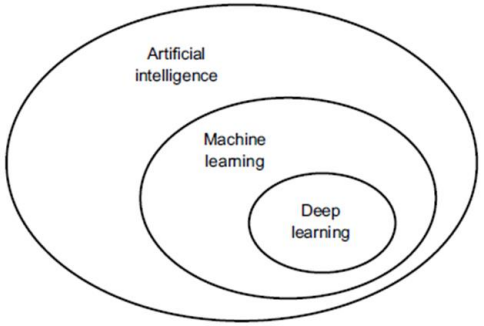
Machine Learning

- A Machine Learning system learns from historical data, builds the prediction models, and whenever it receives new data, predicts the output for it.
- The accuracy of predicted output depends upon the amount of data, as the huge amount of data helps to build a better model which predicts the output more accurately.
- Suppose we have a complex problem, where we need to perform some predictions
- instead of writing a code for it, we need to feed the data to generic algorithms, and with the help of these algorithms, machine builds the logic as per the data and predict the output.



Artificial intelligence

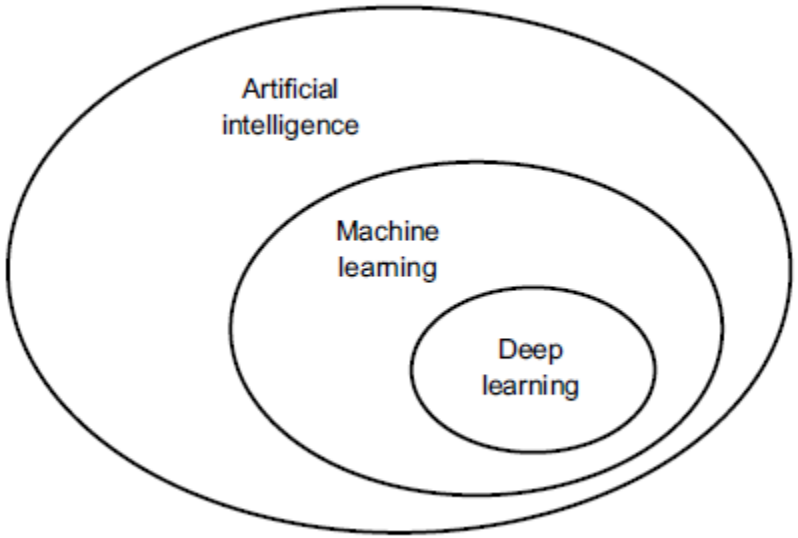
- Concisely, AI can be described as the effort to automate intellectual tasks normally performed by humans.
- AI is a general field that encompasses machine learning and deep learning
- AI also includes many more approaches that may not involve any learning.



Artificial intelligence

- Symbolic AI: Dominant paradigm in AI from the 1950s -1980
- Reached peak popularity during the expert systems boom of the 1980s
- Early chess programs, only involved hardcoded rules crafted by programmers and did not qualify as machine learning.
- For a fairly long time, most experts believed that human-level artificial intelligence could be achieved by having programmers handcraft a sufficiently large set of explicit rules for manipulating knowledge stored in explicit databases.
- This approach is known as symbolic AI.

Artificial intelligence, machine learning, and deep learning



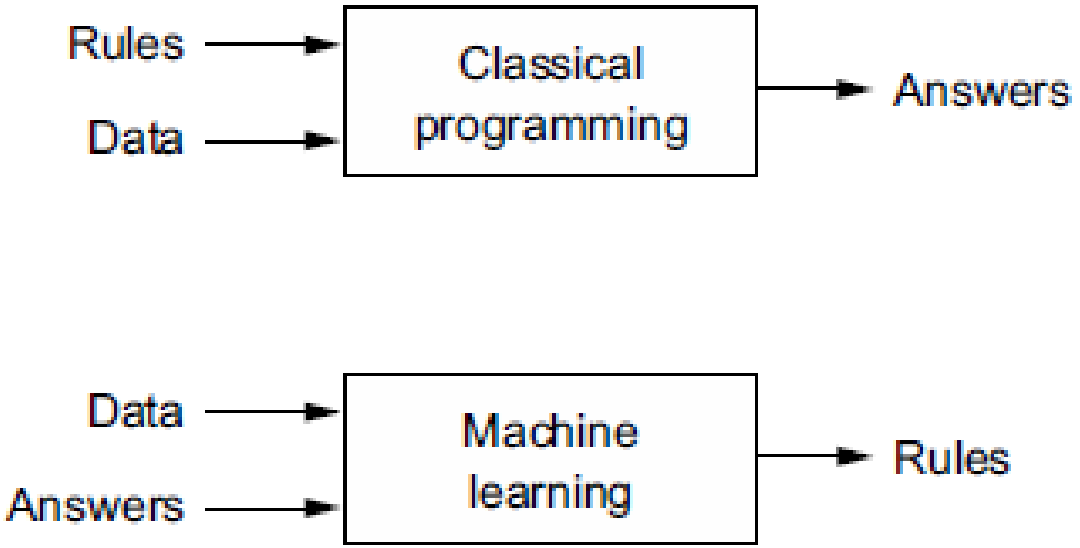
Artificial intelligence

- Symbolic AI to Machine Learning
- Symbolic AI proved suitable to solve well-defined, logical problem
  - such as playing chess
- Turned out to be intractable to figure out explicit rules for solving more complex, fuzzy problem
  - image classification, speech recognition, or natural language translation.
- A new approach arose to take symbolic AI's place: machine learning

Machine Learning

- Machine learning:
- Machine looks at the input data and the corresponding answers, and figures out what the rules should be
- A machine learning system is trained rather than explicitly programmed.
- It is presented with many examples relevant to a task, and it finds statistical structure in these examples that eventually allows the system to come up with rules for automating the task.

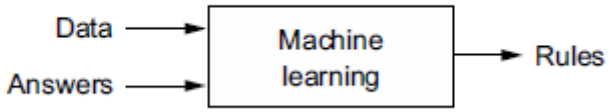
Machine Learning: a new programming paradigm



Machine Learning

- Unlike statistics, machine learning tends to deal with large, complex datasets (such as a dataset of millions of images, each consisting of tens of thousands of pixels)
- Classical statistical analysis such as Bayesian analysis would be impractical.
- machine learning, and especially deep learning, exhibits comparatively little mathematical theory and is fundamentally an engineering discipline

Machine Learning



To do machine learning, we need three things:

1. Input data points—
   - For instance, if the task is speech recognition, these data points could be sound files of people speaking.
   - If the task is image tagging, they could be pictures.
2. Examples of the expected output—
   - In a speech-recognition task, these could be human-generated transcripts of sound files.
   - In an image task, expected outputs could be tags such as “dog,” “cat,” and so on.
3. A way to measure whether the algorithm is doing a good job—

This is necessary to determine the distance between the algorithm’s current output and its expected output.

The measurement is used as a feedback signal to adjust the way the algorithm works.

This adjustment step is what we call learning.

Motivation

- Over the past two decades Machine Learning has become one of the mainstays of information technology
- With the ever increasing amounts of data becoming available, there is good reason to believe that smart data analysis will become more pervasive as a necessary ingredient for technological progress

Applications and type of data

Automatic translation of documents

- At one extreme, we could aim at fully understanding a text before translating it
- using a curated set of rules crafted by a computational linguist well versed in the two languages
- Arduous task
  - text is not always grammatically correct,
  - document understanding is not a trivial task
- Instead, we could use examples of translated documents, such as multilingual entities (United Nations, European Union, Switzerland) to learn how to translate between the two languages
- We could use examples of translations to learn how to translate.
- This machine learning approach proved quite successful

Introduction

- The term machine learning refers to the automated detection of meaningful patterns in data.
- In the past couple of decades it has become a common tool in almost any task that requires information extraction from large data sets.
- We are surrounded by a machine learning based technology:
  - search engines learn how to bring us the best results (while placing profitable ads),
  - anti-spam software learns to filter our email messages, and
  - credit card transactions are secured by a software that learns how to detect frauds.
  - Digital cameras learn to detect faces and
  - intelligent personal assistance applications on smart-phones learn to recognize voice commands
  - Cars are equipped with accident prevention systems that are built using machine learning algorithms.
  - Machine learning is also widely used in scientific applications such as bioinformatics, medicine, and astronomy

Introduction

Common feature of all of these applications

- in contrast to more traditional uses of computers, due to the complexity of the patterns that need to be detected, a human programmer cannot provide an explicit, fine detailed specification of how such tasks should be executed.
- Ex: we human beings, many of our skills are acquired or refined through learning from our experience (rather than following explicit instructions given to us)
- Machine learning tools are concerned with endowing programs with the ability to learn and adapt.

Introduction

- Automated learning, we more often call, Machine Learning (ML)
- That is, we wish to program computers so that they can “learn" from input available to them.
- Roughly speaking, learning is the process of converting experience into expertise or knowledge.
  - The input to a learning algorithm is training data, representing experience, and
  - the output is some expertise, which usually takes the form of another computer program that can perform some task.
- Seeking a formal-mathematical understanding of this concept, we'll have to be more explicit about what we mean by each of the involved terms:
  - What is the training data our programs will access?
  - How can the process of learning be automated?
  - How can we evaluate the success of such a process (namely, the quality of the output of a learning program)?

When Do We Need Machine Learning?

- When do we need machine learning rather than directly program our computers to carry out the task at hand
- Two aspects of a given problem may call for the use of programs that learn and improve on the basis of their “experience":

1. The problem's complexity and
2. the need for adaptivity

When Do We Need Machine Learning?

1. Tasks That Are Too Complex to Program

A. Tasks Performed by Humans: There are numerous tasks that we perform routinely

- yet we cannot sufficiently elaborate to extract a well defined program.
- Ex: driving, speech recognition, and image understanding.

- In all of these tasks, state of the art machine learning programs, programs that “learn from their experience," achieve quite satisfactory results, once exposed to sufficiently many training examples.

When Do We Need Machine Learning?

1. Tasks That Are Too Complex to Program

B. Tasks beyond Human Capabilities:

- Another wide family of tasks that benefit from machine learning techniques are related to the analysis of very large and complex data sets:
  - astronomical data, turning medical archives into medical knowledge, weather prediction, analysis of genomic data, Web search engines, and electronic commerce.
- With more and more available digitally recorded data, it becomes obvious that there are treasures of meaningful information buried in data archives that are way too large and too complex for humans to make sense of.
- Learning to detect meaningful patterns in large and complex data sets is a promising domain in which the combination of programs that learn with the almost unlimited memory capacity and ever increasing processing speed of computers opens up new horizons.

When Do We Need Machine Learning?

2. Adaptivity

- One limiting feature of programmed tools is their rigidity - once the program has been written down and installed, it stays unchanged.
- However, many tasks change over time or from one user to another.
- Machine learning tools - programs whose behavior adapts to their input data - offer a solution to such issues;
- they are, by nature, adaptive to changes in the environment they interact with.
- Typical successful applications of machine learning to such problems include
  - programs that decode handwritten text, where a fixed program can adapt to variations between the handwriting of different users;
  - spam detection programs, adapting automatically to changes in the nature of spam e-mails; and
  - speech recognition programs

Relation to other fields

- As an interdisciplinary field machine learning shares common threads
- With mathematical fields of statistics, information theory, game theory, and optimization
- It is naturally a subfield of computer science
  - as our goal is to program machines so that they will learn.
- machine learning can be viewed as a branch of AI (Artificial Intelligence),
  - the ability to turn experience into expertise or to detect meaningful patterns in complex sensory data is a cornerstone of human intelligence

Relation to other fields

- ML and Statistics
- The component of experience, or training, in machine learning often refers to data that is randomly generated.
- The task of the learner is to process such randomly generated examples toward drawing conclusions that hold for the environment from which these examples are picked.
- This description of machine learning highlights its close relationship with statistics

Relation to other fields

ML and Statistics: Similarities

- ML and Statistics: There is a lot in common between the two disciplines, in terms of both the goals and techniques used.
- if a doctor comes up with the hypothesis that there is a correlation between smoking and heart disease
  - it is the statistician's role to view samples of patients and check the validity of that hypothesis (this is the common statistical task of hypothesis testing).
- In contrast, machine learning aims to use the data gathered from samples of patients to come up with a description of the causes of heart disease.
  - The hope is that automated techniques may be able to figure out meaningful patterns (or hypotheses) that may have been missed by the human observer.

Relation to other fields

- ML vs Statistics: Differences
- In contrast with traditional statistics, in machine learning algorithmic considerations play a major role.
- Machine learning is about the execution of learning by computers;
  - hence algorithmic issues are pivotal.

1. We develop algorithms to perform the learning tasks and are concerned with their computational efficiency.

2. Another difference is that while statistics is often interested in asymptotic behavior (like the convergence of sample-based statistical estimates as the sample sizes grow to infinity),

- theory of machine learning focuses on finite sample bounds.
- Namely, given the size of available samples, machine learning theory aims to figure out the degree of accuracy that a learner can expect on the basis of such samples.

Relation to other fields

ML vs Statistics: Differences

3. In statistics it is common to work under the assumption of certain presubscribed data models (such as assuming the normality of data-generating distributions, or the linearity of functional dependencies)

- in machine learning the emphasis is on working under a “distribution-free" setting, where the learner assumes as little as possible about the nature of the data distribution and allows the learning algorithm to figure out which models best approximate the data-generating process.

First ML Applications?

- ML has been around in some specialized applications, such as Optical Character Recognition (OCR).
- First ML application that really became mainstream in in the 1990s: the spam filter.
- It does technically qualify as Machine Learning (actually learned so well that we seldom need to flag an email as spam anymore).
- It was followed by hundreds of ML applications that now quietly power hundreds of products and features that we use regularly
- from better recommendations to voice search

What Is Machine Learning? – Informal Definition

- Machine learning can be defined as the process of solving a practical problem by

1) gathering a dataset, and
2) algorithmically building a statistical model based on that dataset

- That statistical model is assumed to be used to solve the practical problem.

What Is Machine Learning? – Formal Definition

- Machine Learning is the science (and art) of programming computers so they can learn from data.
- Definitions:
- [Machine Learning is the] field of study that gives computers the ability to learn without being explicitly programmed. —Arthur Samuel, 1959
- Engineering-oriented: A computer program is said to learn from experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E. —Tom Mitchell, 1997

WELL-POSED MACHINE LEARNING PROBLEMS

In general, to have a well-defined learning problem, we must identify these three features:

- The learning task
- The measure of performance
- The task experience

- The computer program is the ‘machine’ in our context.
- The computer program is designed employing learning from the task experience.
- Equivalently, we say that the machine is trained using task experience, or machine learns from task experience
- Understanding the inputs and outputs is of greater importance

WELL-POSED MACHINE LEARNING PROBLEMS

- The input is defined by the learning task.
- Four different types of learning tasks appear in the real-world applications (supervised, unsupervised, reinforcement, learning based on natural processes)
- For different forms of raw data (text, images, waveforms, and so forth), it is common to represent data in standard fixed length vector formats with numerical values.
- Such abstractions typically involve significant loss of information, yet they are essential for a well-defined learning problem.

WELL-POSED MACHINE LEARNING PROBLEMS

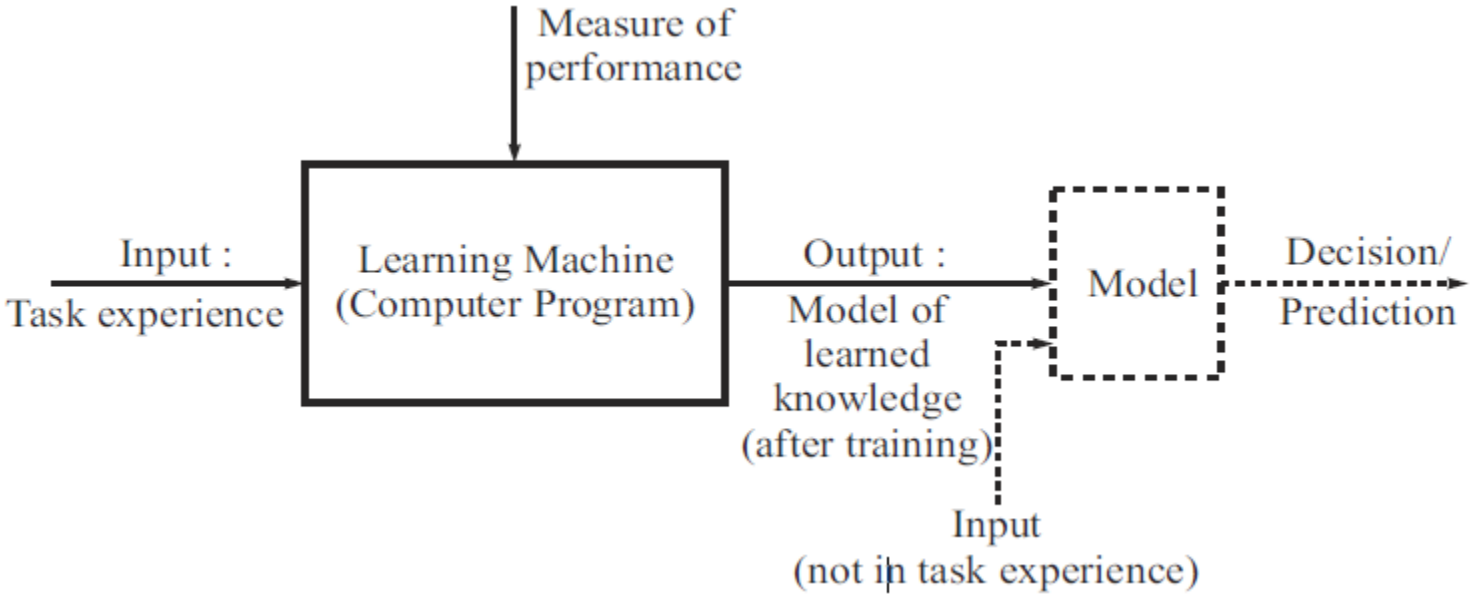
- Numerical form of data representation allows us to deal with patterns geometrically
  - So we use learning algorithms using linear algebra and analytic geometry
  - Characterizing the similarity of patterns in state space can be done through some form of metric (distance) measure:
  - distance between two vectors is a measure of similarity between two corresponding patterns.
  - Many measures of ‘distance’ have been proposed in the literature.
- In another class of machine leaning problems, the input (experience) is available in the form of
  - nominal (or categorical) data, described in linguistic form (not numerical).
  - For nominal form of data, there is no natural notion of similarity.
  - Each learning algorithm based on nominal data employs some nonmetric method of similarity.

WELL-POSED MACHINE LEARNING PROBLEMS

- The output of an algorithm represents the learned knowledge.
- This knowledge is in the form of a model of the structural patterns in the data.
- The model is deployed by the user for decision-making;
- it gives the prediction with respect to the assigned task for measurements/observations not in the task experience;
- a good model will generalize well to observations unseen by the machine during training.

WELL-POSED MACHINE LEARNING PROBLEMS

A block diagrammatic representation of a learning machine



Spam filter - ML Application

- Spam filter is a Machine Learning program
- Given examples of spam emails (e.g., flagged by users) and examples of regular (nonspam, also called “ham”) emails can learn to flag spam.
- The examples that the system uses to learn are called the training set.
- Each training example is called a training instance (or sample).
- Task T is to flag spam for new emails
- Experience E is the training data
- Performance measure P needs to be defined;
  - We can use the ratio of correctly classified emails.
  - This particular performance measure is called accuracy, and it is often used in classification tasks.

Spam filter Application using traditional programming technique

- What spam typically looks like?

1. Notice that some words or phrases (such as “4U,” “credit card,” “free,” and “amazing”) tend to come up a lot in the subject line.

- Perhaps a few other patterns in the sender’s name, the email’s body, and other parts of the email could also be noticed

2. Write a detection algorithm for each of the patterns noticed, and program would flag emails as spam if a number of these patterns were detected.

- 3. Test our program and repeat steps 1 and 2 until it was good enough

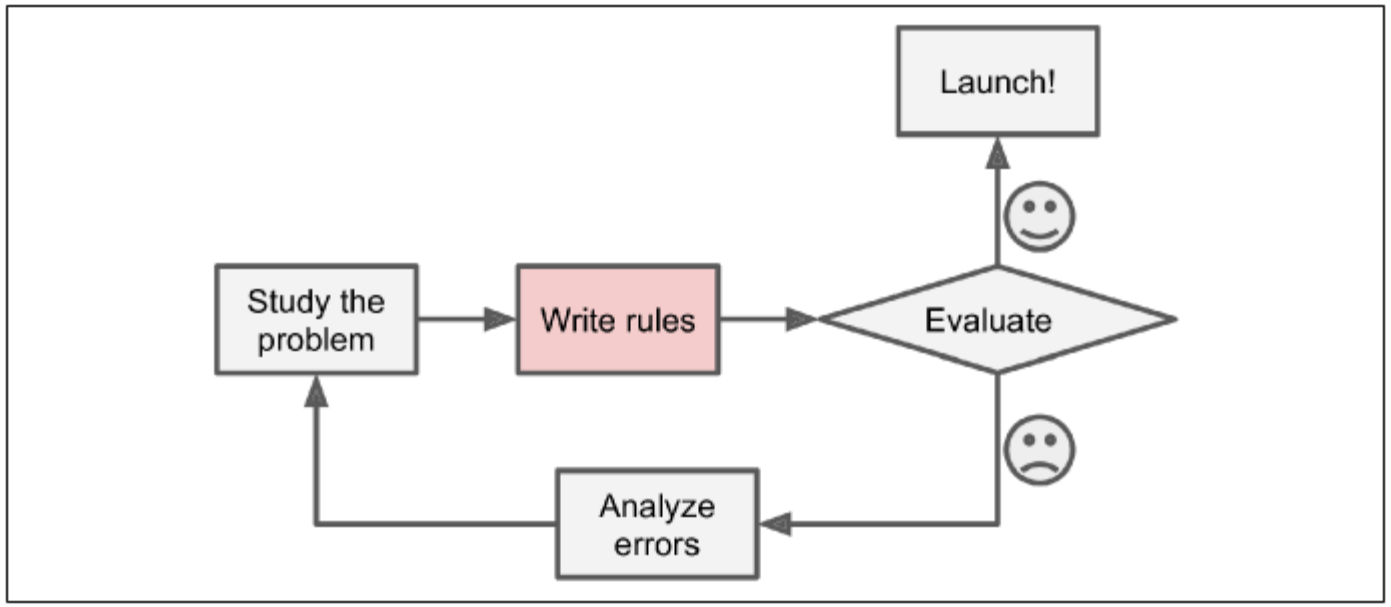
to launch.

Problem?

Program will likely become a long list of complex rules—hard to maintain

Spam filter Application using traditional programming technique

- Write a spam filter using traditional programming techniques

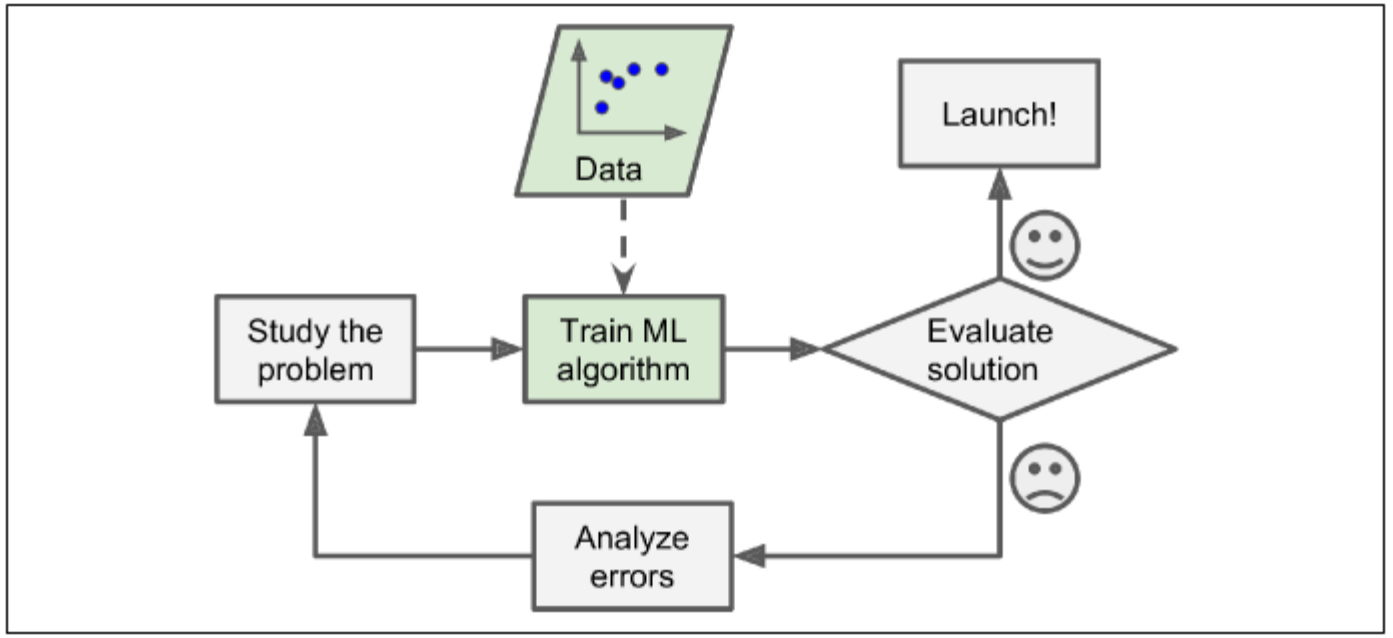


Spam filter using ML Approach

- a spam filter based on Machine Learning techniques automatically learns which words and phrases are good predictors of spam
- by detecting unusually frequent patterns of words in the spam examples compared to the ham examples
- The program is much shorter, easier to maintain, and most likely more accurate.

Spam filter using ML Approach

- Write a spam filter using ML Approach

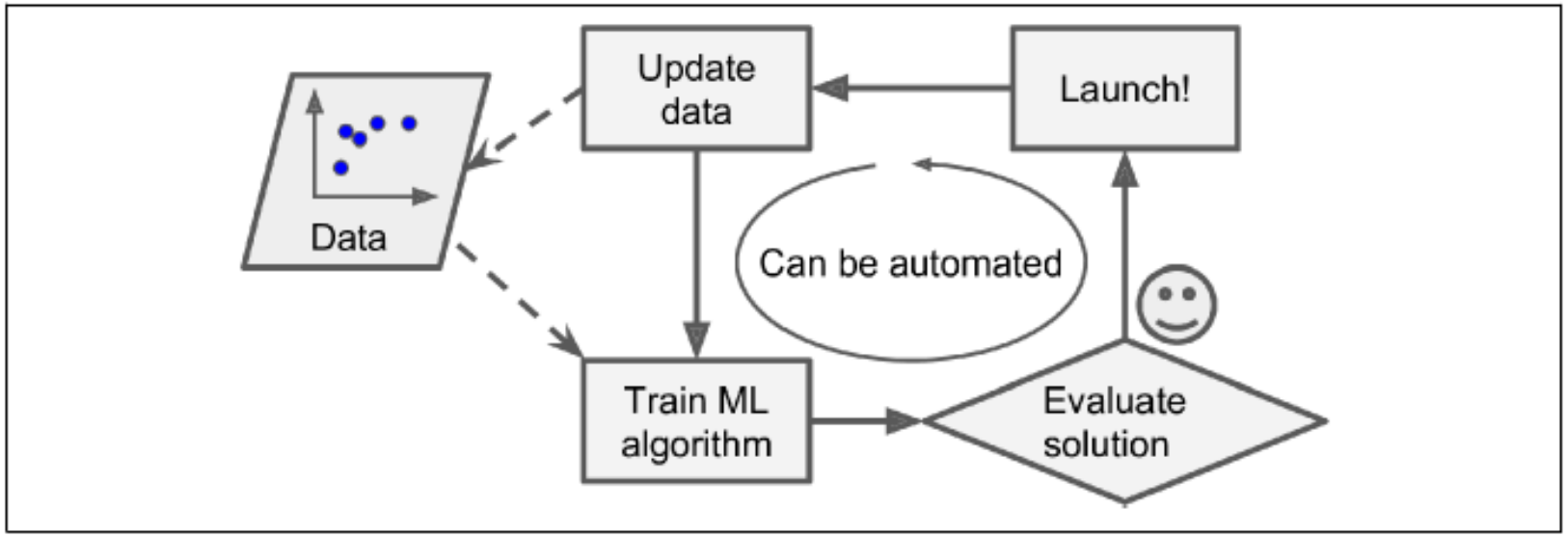


Spam filter using ML Approach

- Automatically adapting to change
- If spammers keep working around our spam filter, we need to keep writing new rules forever
- Ex: What if spammers notice that all their emails containing “4U” are blocked?
- They might start writing “For U” instead.
- A spam filter using traditional programming techniques would need to be updated to flag “For U” emails.
- In contrast, a spam filter based on Machine Learning techniques automatically notices that “For U” has become unusually frequent in spam flagged by users, and it starts flagging them without our intervention

Spam filter using ML Approach

- Write a spam filter using ML Approach- Automatically adapting to change



Other Applications using ML Approach

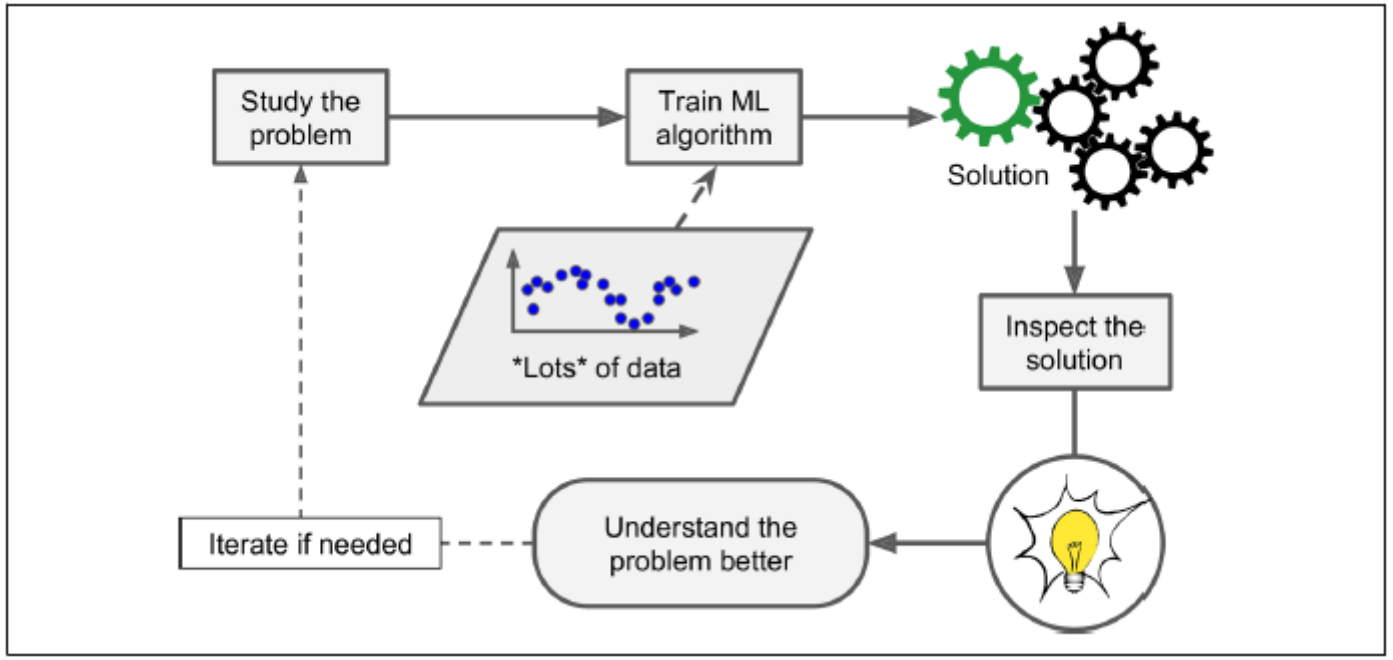
- Another area where Machine Learning shines is for problems that either
  - are too complex for traditional approaches or
  - have no known algorithm.
- Ex: Speech recognition
- To write a program capable of distinguishing the words “one” and “two.”
- We might notice that the word “two” starts with a high-pitch sound (“T”)
- so we could hardcode an algorithm that measures high-pitch sound intensity and use that to distinguish ones and twos
- Problem?
- But this technique will not scale to thousands of words spoken by millions of very different people in noisy environments and in dozens of languages.
- The best solution is to write an algorithm that learns by itself, given many example recordings for each word.

Other Applications using ML Approach

- Machine Learning can help humans learn
- ML algorithms can be inspected to see what they have learned (although for some algorithms this can be tricky).
- Ex: once a spam filter has been trained on enough spam, it can easily be inspected to reveal
  - the list of words and
  - combinations of words that it believes are the best predictors of spam.
- Sometimes this will reveal unsuspected correlations or new trends, and thereby lead to a better understanding of the problem.
- Applying ML techniques to dig into large amounts of data can help discover patterns that were not immediately apparent.
- This is called data mining

Other Applications using ML Approach

Machine Learning can help humans learn



Why Use Machine Learning?- Summary

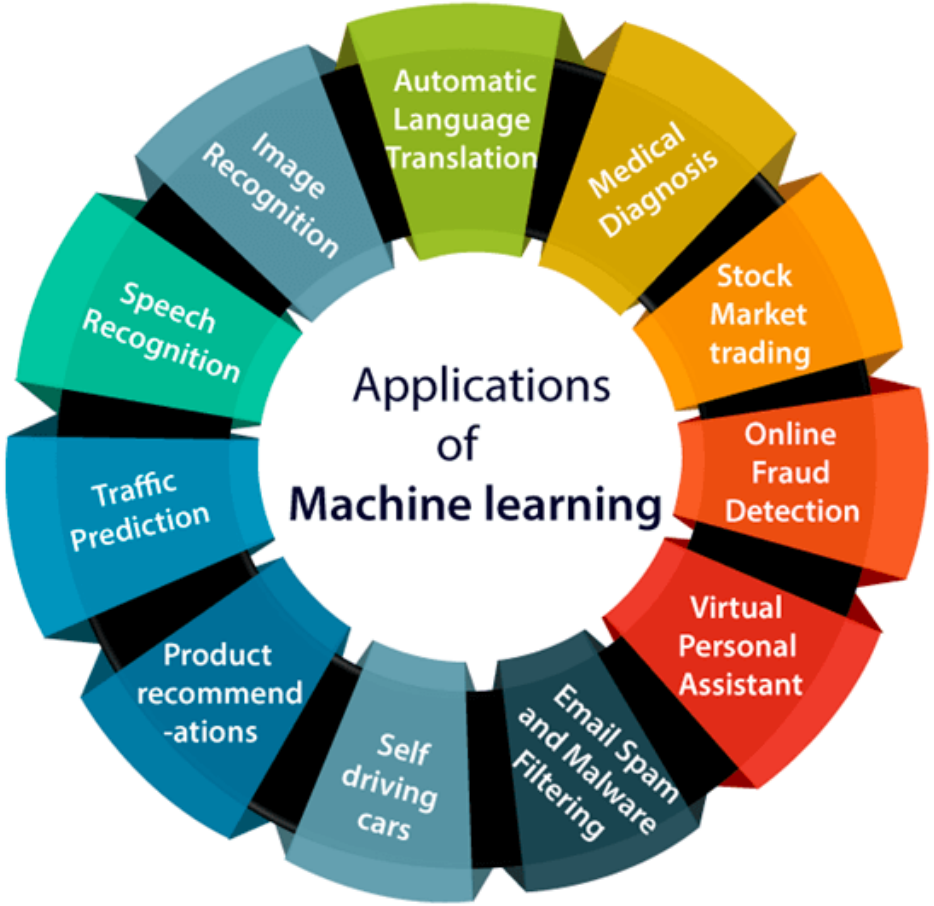
Machine Learning is great for:

- Problems for which existing solutions require a lot of fine-tuning or long lists of rules:
  - one Machine Learning algorithm can often simplify code and perform better than the traditional approach.
- Complex problems for which using a traditional approach yields no good solution:
  - the best Machine Learning techniques can perhaps find a solution.
- Fluctuating environments: a Machine Learning system can adapt to new data.
- Getting insights about complex problems and large amounts of data.

EXAMPLES OF APPLICATIONS IN DIVERSE FIELDS

- Machine learning is a growing technology used to mine knowledge from data (popularly known as data mining field).
- Wherever data exists, things can be learned from it.
  - Whenever, there is excess of data, the mechanics of learning must be automatic.
  - Machine learning technology is meant for automatic learning from voluminous datasets.
- Applications emerge not from machine learning experts, nor from the data itself, but from people who work with the data and the problems from which it arises.

EXAMPLES OF APPLICATIONS IN DIVERSE FIELDS



If we download a copy of Wikipedia, has my computer really learned something? Is it suddenly smarter?

- If we just download a copy of Wikipedia, our computer has a lot more data,
- but it is not suddenly better at any task. Thus, downloading a copy of Wikipedia is not Machine Learning.

ML Future

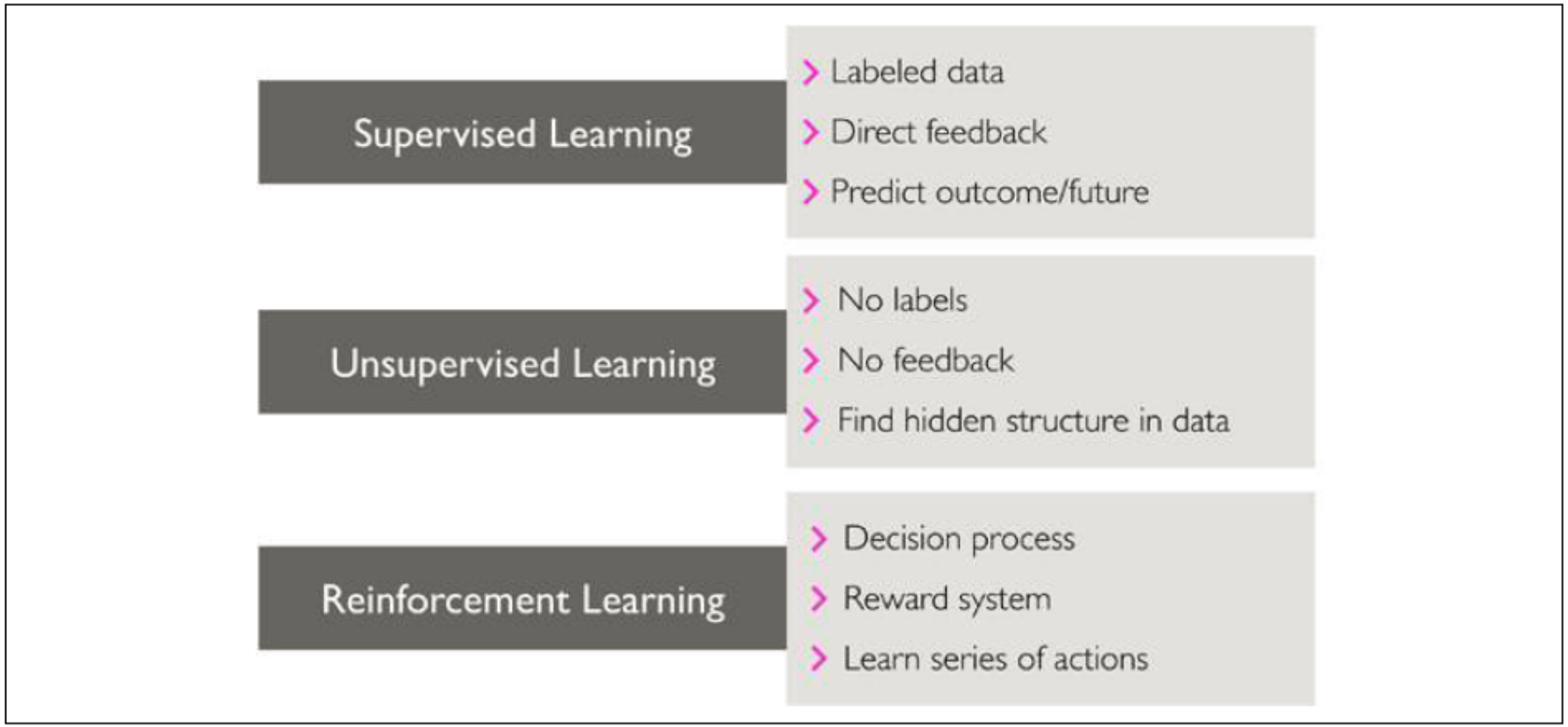
- Hopefully soon there will be safe and efficient self-driving cars
- Notable progress has been made in medical applications;
  - researchers demonstrated that deep learning models can detect skin cancer with near-human accuracy (https://www.nature.com/articles/nature21056).
- Another milestone was recently achieved by researchers
  - DeepMind, who used deep learning to predict 3D protein structures, outperforming physics-based approaches for the first time (https://deepmind.com/blog/alphafold/).

FORMS OF LEARNING

- In the broadest sense, any method that incorporates information from experience in the design of a machine, employs learning.
- A learning method depends on the type of experience from which the machine will learn (with which the machine will be trained).
- The type of available learning experience can have significant impact on success or failure of the learning machine.
- The field of machine learning usually distinguishes four forms of learning:

1. supervised learning,
2. unsupervised learning,
3. reinforcement learning,
4. learning based on natural processes—evolution, swarming, and immune systems

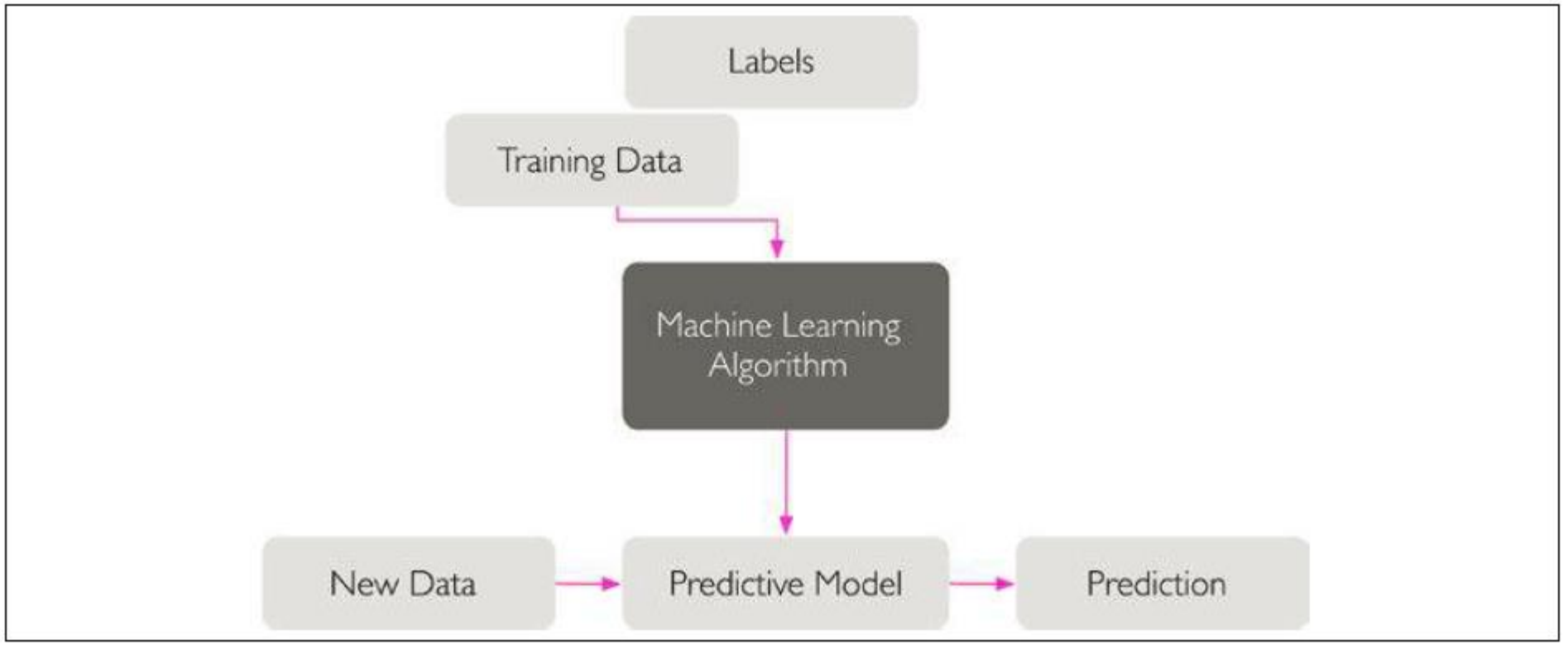
FORMS OF LEARNING



FORMS OF LEARNING – Supervised Learning

A supervised learning workflow

- the labeled training data is passed to a machine learning algorithm for fitting a predictive model that can make predictions on new, unlabeled data inputs:

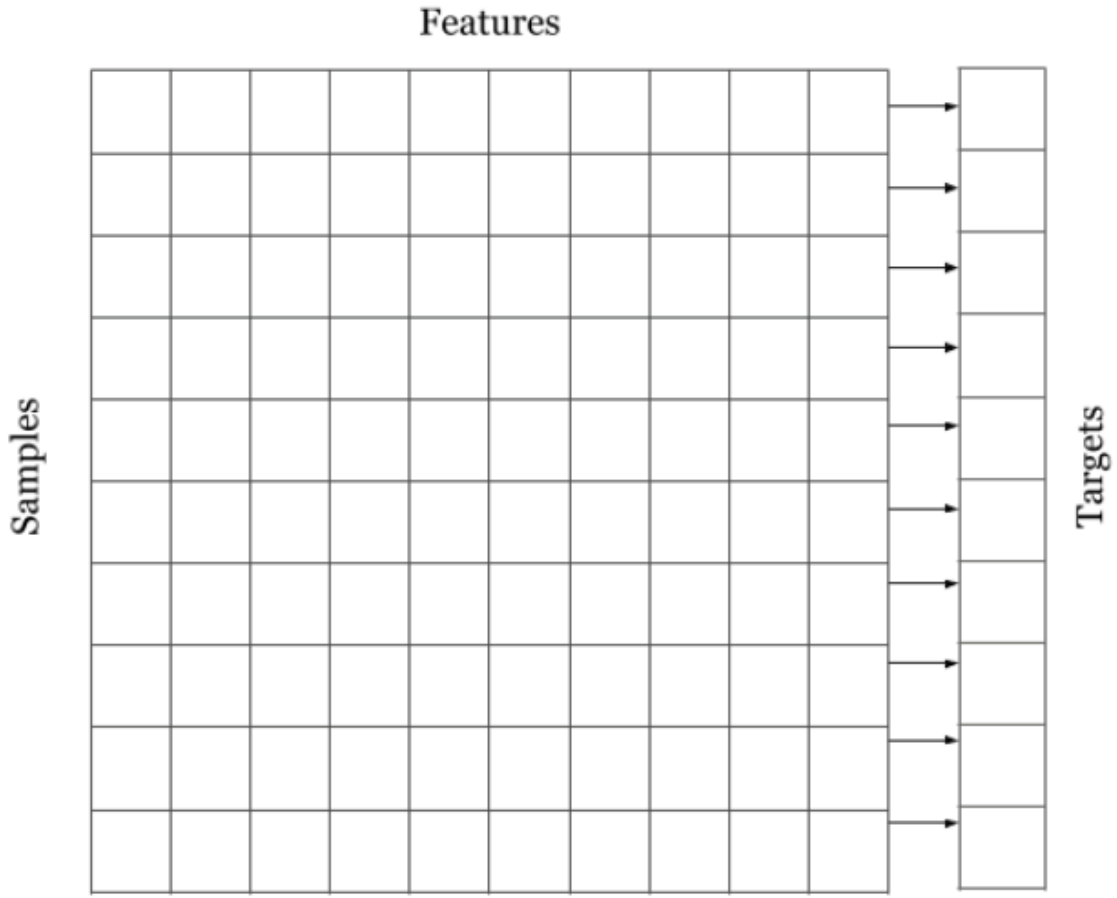


FORMS OF LEARNING – Supervised Learning

- The main goal in supervised learning is to learn a model from labeled training data that allows us to make predictions about unseen or future data.
- Here, the term "supervised" refers to a set of training examples (data inputs) where the desired output signals (labels) are already known.

FORMS OF LEARNING – Supervised Learning

A supervised dataset



FORMS OF LEARNING – Supervised Learning

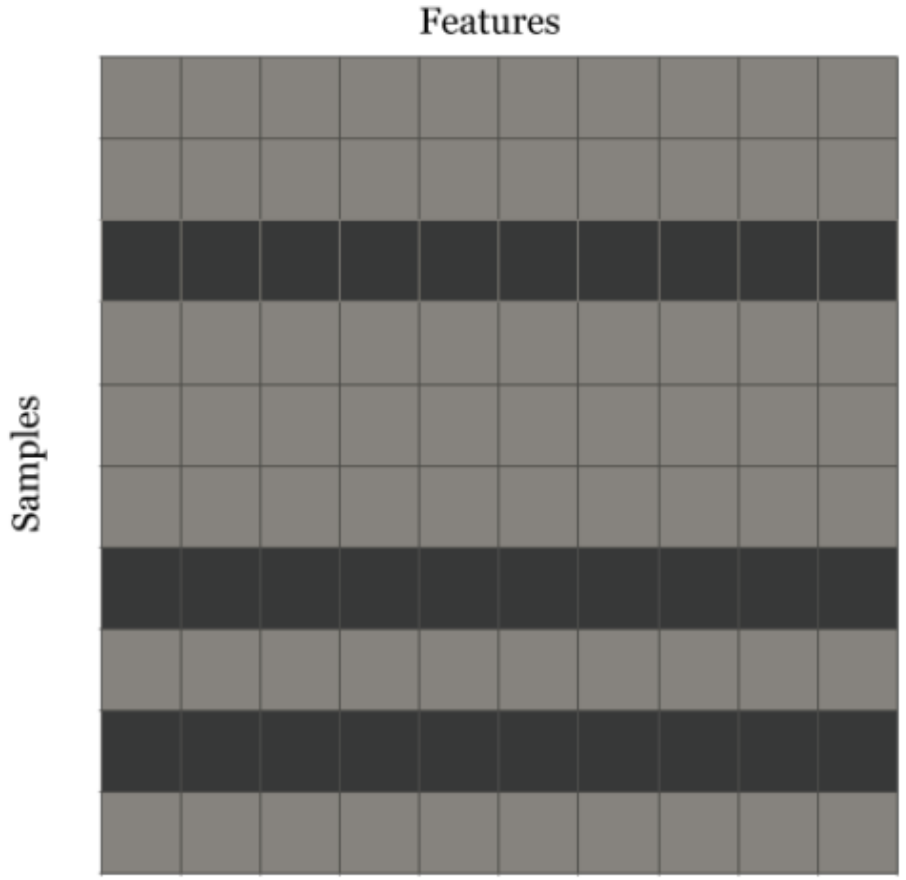
- A problem in which we are required to predict a value is known as a supervised problem.
- Ex1:
- if the problem is to predict house prices given historical house prices, with features like presence of a hospital, school or supermarket, distance to nearest public transport, etc. is a supervised problem.
- Ex2:
- Similarly, when we are provided with images of cats and dogs, and we know beforehand which ones are cats and which ones are dogs, and if the task is to create a model which predicts whether a provided image is of a cat or a dog, the problem is considered to be supervised.

FORMS OF LEARNING – Supervised Learning

- In figure, every row of the data is associated with a target or label.
- The columns are different features and rows represent different data points which are usually called samples.
- The example shows ten samples with ten features and a target variable which can be either a number or a category.
- If the target is categorical, the problem becomes a classification problem.
- If the target is a real number, the problem is defined as a regression problem.
- Supervised problems can be divided into two sub-classes:
  - Classification: predicting a category, e.g. dog or cat.
  - Regression: predicting a value, e.g. house prices.

FORMS OF LEARNING – Unsupervised Learning

An unsupervised dataset



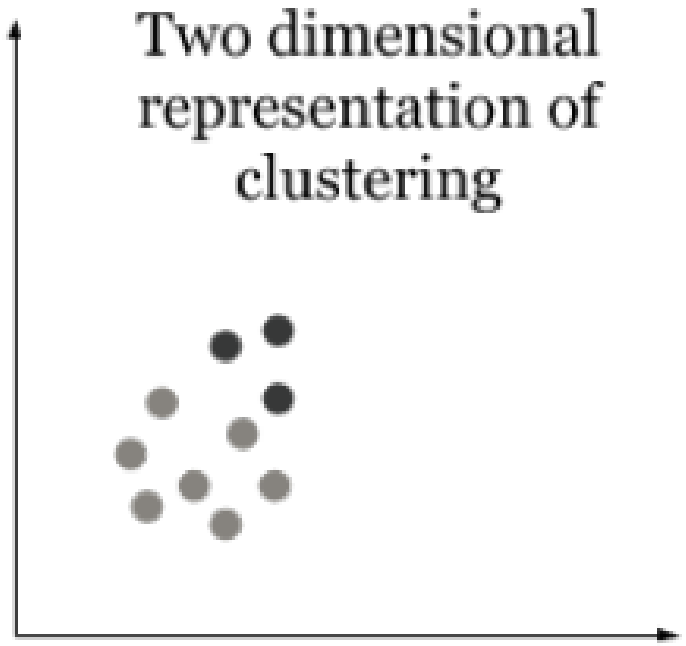
FORMS OF LEARNING – Unsupervised Learning

- Unsupervised datasets do not have a target associated with them
- Ex1:
- A financial firm deals with credit card transactions.
- There is a lot of data that comes in every second.
- The only problem is that it is difficult to find humans who will mark each and every transaction either as a valid or genuine transaction or a fraud.
- When we do not have any information about a transaction being fraud or genuine, the problem becomes an unsupervised problem.
- To tackle these kinds of problems we have to think about how many clusters can data be divided into.
- Clustering is one of the approaches for problems but there are several other approaches
- For a fraud detection problem, we can say that data can be divided into two classes (fraud or genuine).

FORMS OF LEARNING – Unsupervised Learning

- When we know the number of clusters, we can use a clustering algorithm for unsupervised problems.
- In figure, the data is assumed to have two classes,
  - Dark colour represents fraud, and
  - light colour represents genuine transactions.
- These classes, however, are not known to us before the clustering approach.
- After a clustering algorithm is applied, we should be able to distinguish between the two assumed targets.
- To make sense of unsupervised problems, we can also use numerous decomposition techniques such as Principal Component Analysis
- (PCA), t-distributed Stochastic Neighbour Embedding (t-SNE) etc.

FORMS OF LEARNING – Unsupervised Learning



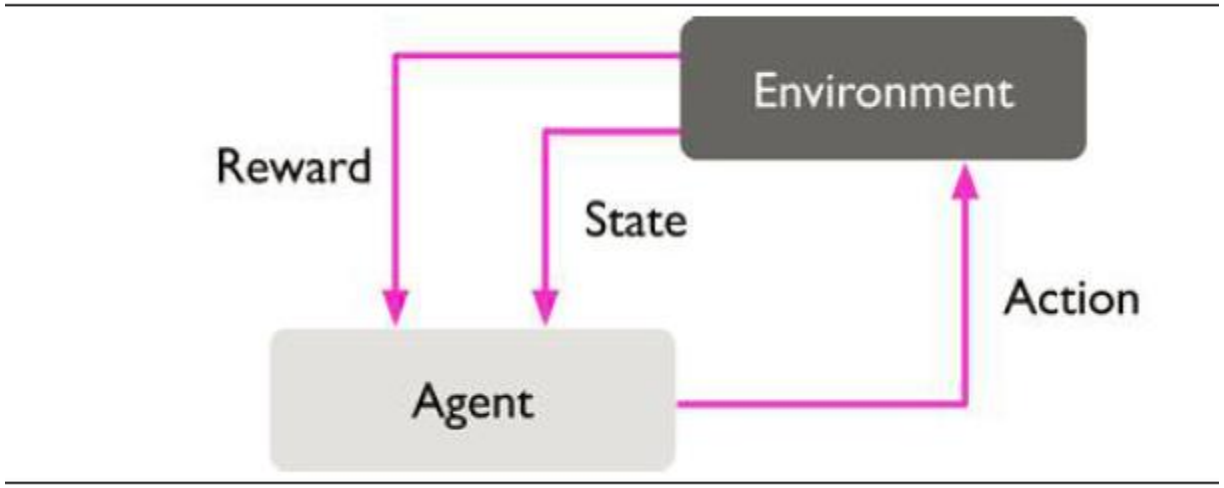
FORMS OF LEARNING – Reinforcement Learning

- Reinforcement learning is founded on the concept that if an action is followed by a satisfactory state of affairs, or by an improved state of affairs (according to some properly defined way), then the inclination to produce that action becomes stronger, i.e., reinforced.
- This idea can be extended to permit action choices to be dependent on state information, which then brings in the aspect of feedback.
- A reinforcement learning system, therefore, is a system that via interaction with its environment enhances its performance by obtaining feedback in the form of a scalar reward (or penalty)—a reinforcement signal, that is indicative of the suitability of the response.
- The learning system is not instructed with regard to what action has to be taken.
- Instead, it is expected to find out which actions produce the maximum reward by trying them.
- The actions may influence not only the immediate reward but also the next situation, and through that all subsequent rewards

Reinforcement Learning

- Reinforcement learning is a feedback-based learning method, in which a learning agent gets a reward for each right action and gets a penalty for each wrong action.
- The agent learns automatically with these feedbacks and improves its performance.
- In reinforcement learning, the agent interacts with the environment and explores it.
- The goal of an agent is to get the most reward points, and hence, it improves its performance.
- The robotic dog, which automatically learns the movement of his arms, is an example of Reinforcement learning.

Reinforcement Learning



Reinforcement Learning- Summary

- In reinforcement learning, the goal is to develop a system (agent) that improves its performance based on interactions with the environment.
- Since the information about the current state of the environment typically also includes a so-called reward signal, we can think of reinforcement learning as a field related to supervised learning.
- However, in reinforcement learning, this feedback is not the correct ground truth label or value, but a measure of how well the action was measured by a reward function.
- Through its interaction with the environment, an agent can then use reinforcement learning to learn a series of actions that maximizes this reward via an exploratory trial-and-error approach or deliberative planning.

FORMS OF LEARNING - Learning Based on Natural Processes: Evolution, Swarming, and Immune Systems

- Some learning approaches take inspiration from nature for the development of novel problem solving techniques.
- The thread that ties together learning based on evolution process, swarm intelligence, and immune systems is that all have been applied successfully to a variety of optimization problems.
- Optimization may not appear to be like a machine learning task, but optimization techniques are commonly used as part of machine learning algorithms.