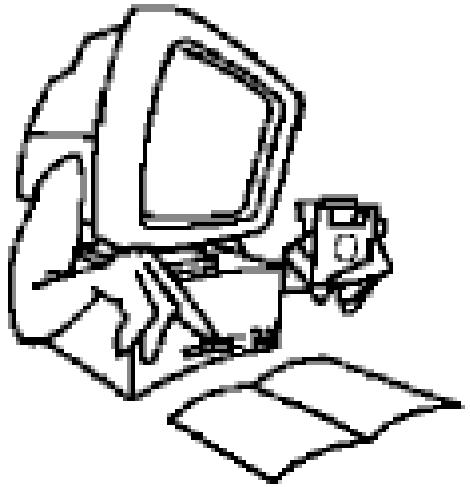


MACHINE LEARNING

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

LEARNING

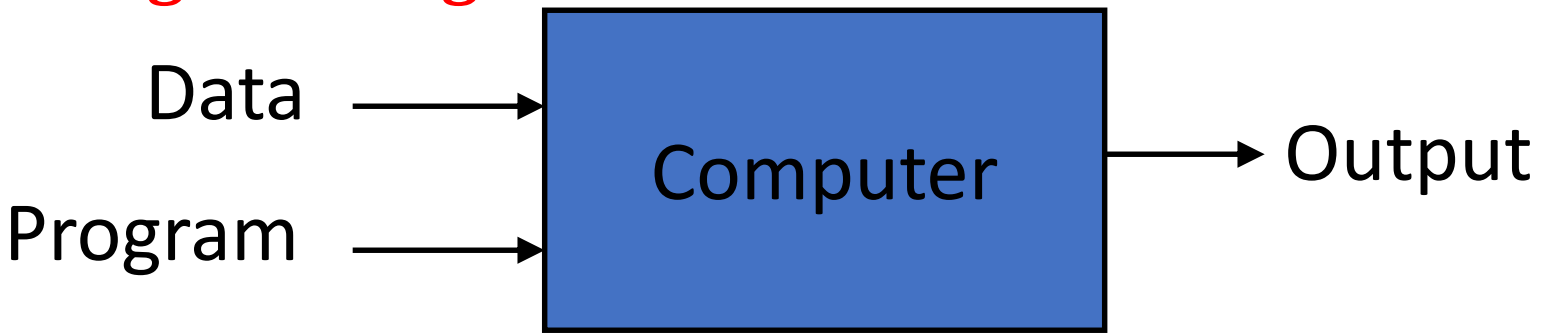
- "Learning denotes changes in a system that enables a system to do the same task more efficiently the next time." – (Herbert Simon)

- "Learning is constructing or modifying representations of what is being experienced." – (Ryszard Michalski)

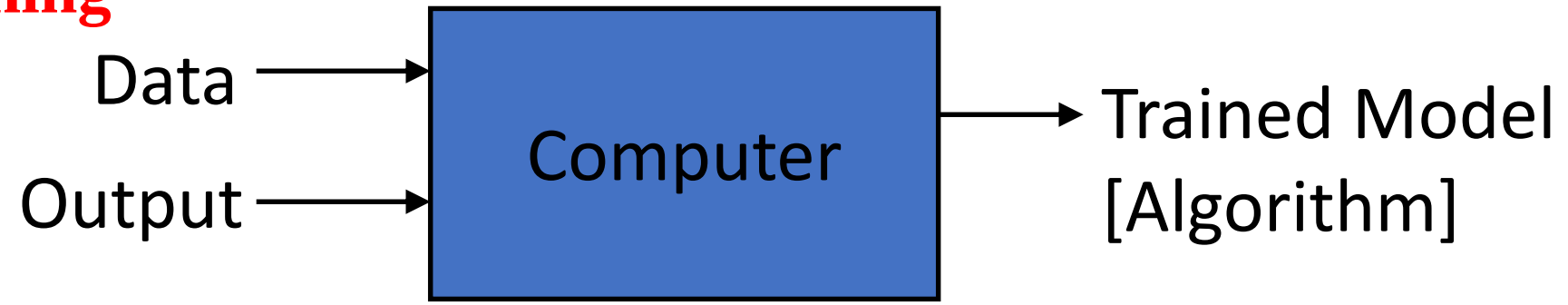
- "Learning is making useful changes in our minds." –(Marvin Minsky)

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Traditional Programming



Machine Learning



Machine Learning Definition

- Field of study that gives the computer the ability to learn without being explicitly programmed. (Arthur Samuel, 1959)
- 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, 1998)

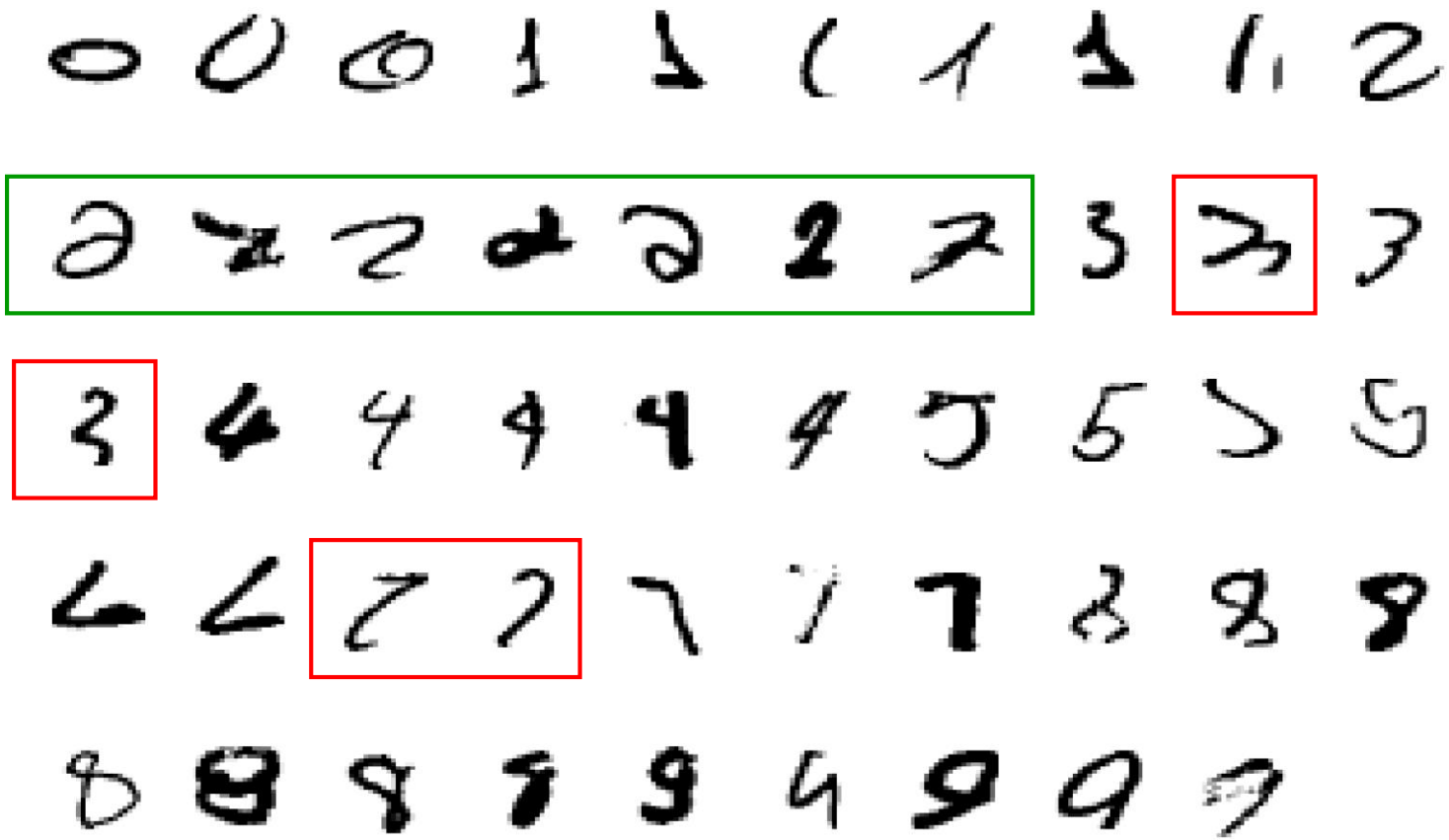
Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Machine Learning (Wiki)

- Machine learning is the subfield of artificial intelligence that is concerned with the design and development of algorithms that allow computers to improve their performance over time based on data, such as from sensor data or databases.
- A major focus of machine learning research is to automatically produce (induce) models, such as rules and patterns, from data.
- Hence, machine learning is closely related to fields such as data mining, statistics, inductive reasoning, pattern recognition, and theoretical computer science.

Why Machine Learning Is Hard, Redux

What is a “2”?



When Would We Use Machine Learning?

- When patterns exists in our data
  - Even if we don't know what they are
    - Or perhaps especially when we don't know what they are
- We can not pin down the functional relationships mathematically
  - Else we would just code up the algorithm
- When we have lots of (unlabeled) data
  - Labeled training sets harder to come by
  - Data is of high-dimension
    - High dimension "features"
    - For example, sensor data
  - Want to "discover" lower-dimension representations
    - Dimension reduction

Learning in AI

- Learning is essential for unknown environments,
  - i.e., when designer lacks omniscience
- Learning is useful as a system construction method,
  - i.e., expose the agent to reality rather than trying to write it down
- Learning modifies the agent's decision mechanisms to improve performance
- Understand and improve efficiency of human learning
  - Use to improve methods for teaching and tutoring people
  - (e.g., better computer-aided instruction.)

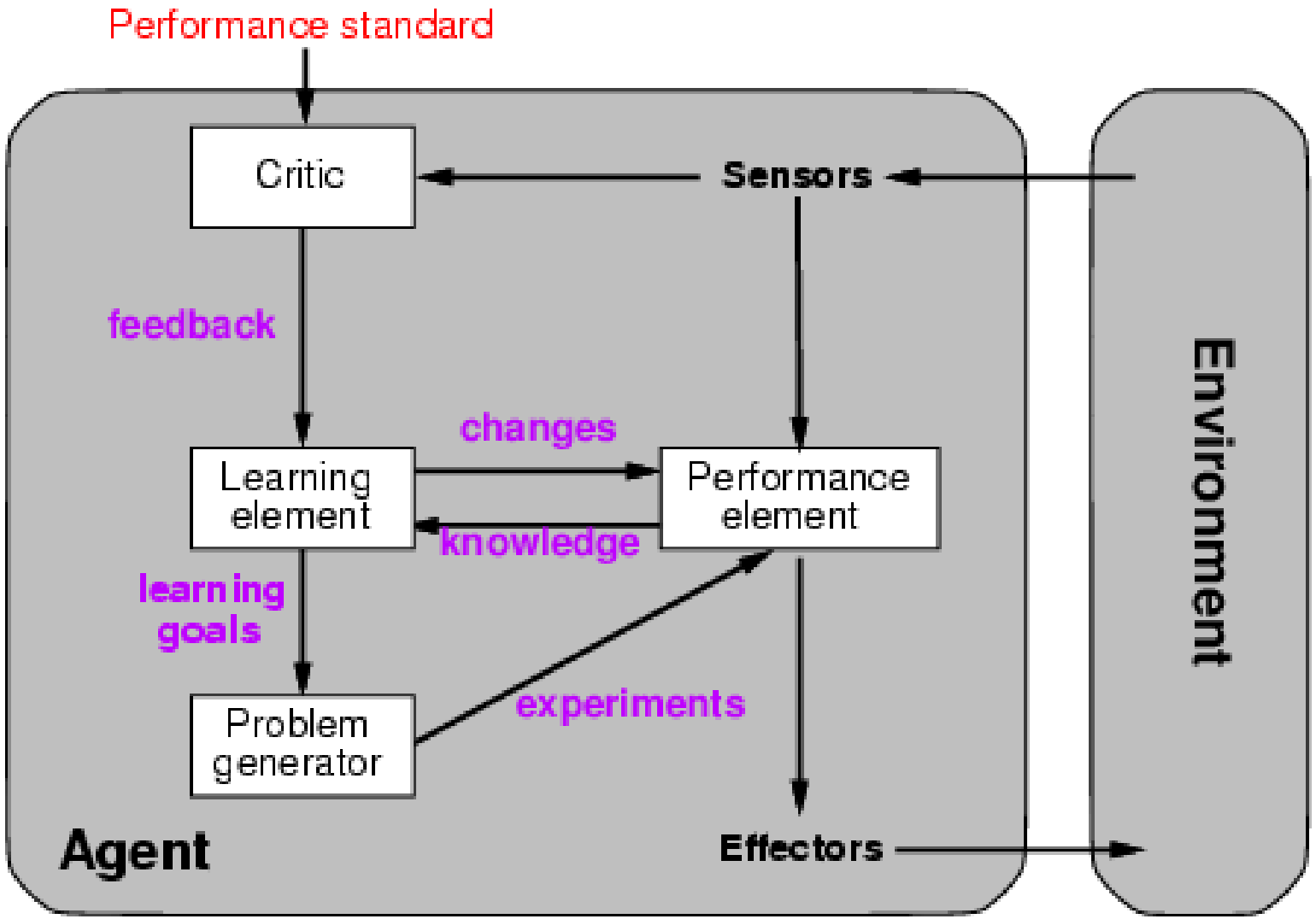
Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Learning in AI

- Discover new things or structure that is unknown to humans
  - e.g., Data mining, Knowledge Discovery in Databases
- Fill in skeletal or incomplete specifications about a domain
  - Large, complex AI systems cannot be completely derived by hand and require dynamic updating to incorporate new information.
  - Learning new characteristics expands the domain or expertise and lessens the "brittleness" of the system
- Build software agents that can adapt to their users or to other software agents.

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Learning agent



Learning element

- Design of a learning element is affected by
  - Which components of the performance element are to be learned
  - What feedback is available to learn these components
  - What representation is used for the components

- Type of feedback:
  - Supervised learning: correct answers for each example
  - Unsupervised learning: correct answers not given
  - Reinforcement learning: occasional rewards

There are three important roles for machine learning.

1. **Data Mining:** this is using historical data to improve decisions. An example is looking at medical records and applying it to medical knowledge when making a diagnoses.
2. **Software applications that we cannot program by hand:** Examples of this are autonomous driving and speech recognition
3. **Self-customizing programs:** An example of this is a newsreader that learns a readers particular interests and highlights these when the reader visits the site.

Machine Learning Tasks

- **Categorisation**
  - Learn why certain objects are categorised a certain way
  - Learn attributes of members of each category from background information.
- **Prediction**
  - Learn how to predict how to categorise unseen objects
  - E.g., given examples of financial stocks and a categorisation of them into safe and unsafe stocks
  - Learn how to predict whether a new stock will be safe

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Potentials for Machine Learning

Agents can learn these from examples:

- which chemicals are toxic (*biochemistry*)
- which patients have a disease (*medicine*)
- which substructures proteins have (*bioinformatics*)
- what the grammar of a language is (*natural language*)
- which stocks and shares are about to drop (*finance*)
- which vehicles are tanks (*military*)
- which style a composition belongs to (*music*)
- And many more…

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Examples of Machine Learning Problems

- Pattern Recognition
  - Facial identities or facial expressions
  - Handwritten or spoken words (e.g., Siri)
  - Medical images
  - Sensor Data/IoT
- Optimization
  - Many parameters have “hidden” relationships that can be the basis of optimization
- Pattern Generation
  - Generating images or motion sequences
- Anomaly Detection
  - Unusual patterns in the telemetry from physical and/or virtual plants (e.g., data centers)
  - Unusual sequences of credit card transactions
  - Unusual patterns of sensor data from a nuclear power plant
    - or unusual sound in your car engine or …
- Prediction
  - Future stock prices or currency exchange rates
  - Network events

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Performing Machine Learning

- Specify your problem as a learning task
- Choose the representation scheme
- Choose the learning method
- Apply the learning method
- Assess the results and the learning method (*evaluation*)

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Supervised Machine Learning

- We are given input samples (X) and output samples (y) of a function

  $y = f(X)$.

- We would like to "learn" f and evaluate it on new data.

- Types:
  - **Classification:** y is discrete (class labels).
  - **Regression:** y is continuous, e.g. linear regression.

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

- **Examples of Supervised Learning:**
  - Is this image a cat, dog, car, house?
  - How would this user score that restaurant?
  - Is this email spam?
  - Is this blob a supernova?

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Usefulness of Supervised Learning

- There are two main areas where supervised learning is useful: classification problems and regression problems.

- **Classification** problems ask the algorithm to predict a discrete value, identifying the input data as the member of a particular class, or group. In a training dataset of animal images, that would mean each photo was pre-labeled as cat, koala or turtle. The algorithm is then evaluated by how accurately it can correctly classify new images of other koalas and turtles.

- On the other hand, **regression** problems look at continuous data. One use case, linear regression, should sound familiar from algebra class: given a particular x value, what's the expected value of the y variable?

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

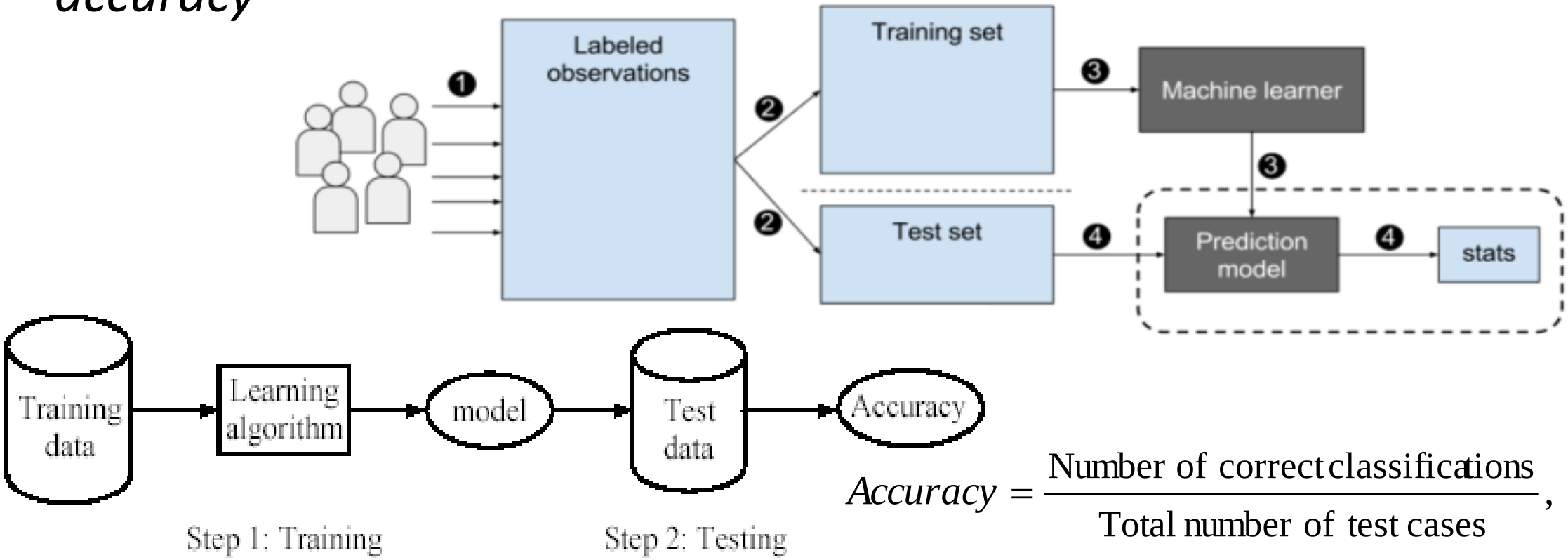
Supervised Learning

- Supervised learning means having a full set of labeled data while training an algorithm.
- Fully labeled means that each example in the training dataset is tagged with the answer the algorithm should come up with on its own.
- A labeled dataset of flower images would tell the model which photos were of roses, daisies and daffodils.
- When shown a new image, the model compares it to the training examples to predict the correct label.

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

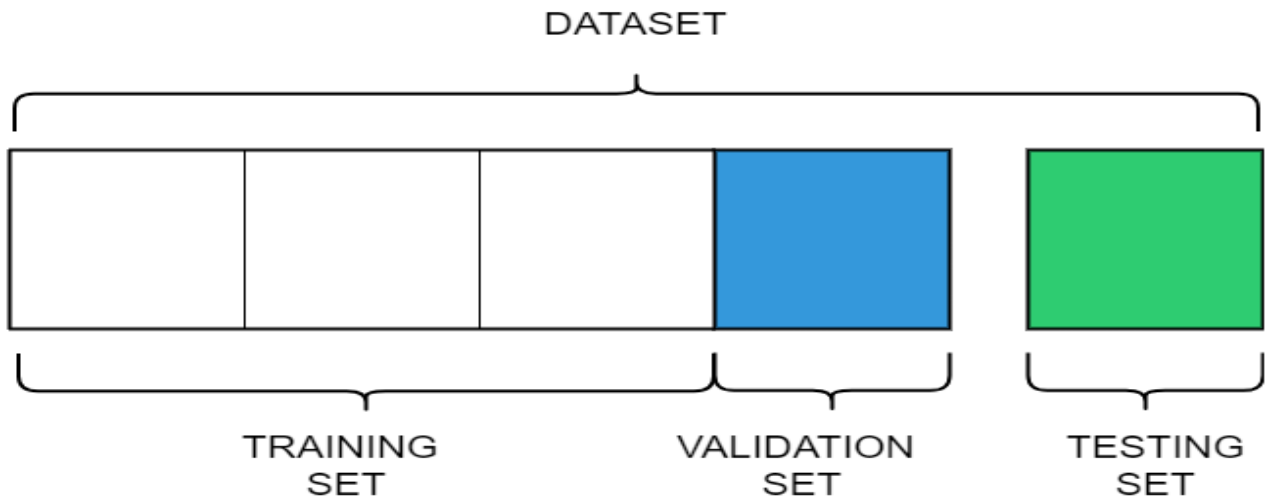
Supervised Learning Process

- Learning (training): *Learn a model using the training data*
- Testing: *Test the model using unseen test data to assess the model accuracy*

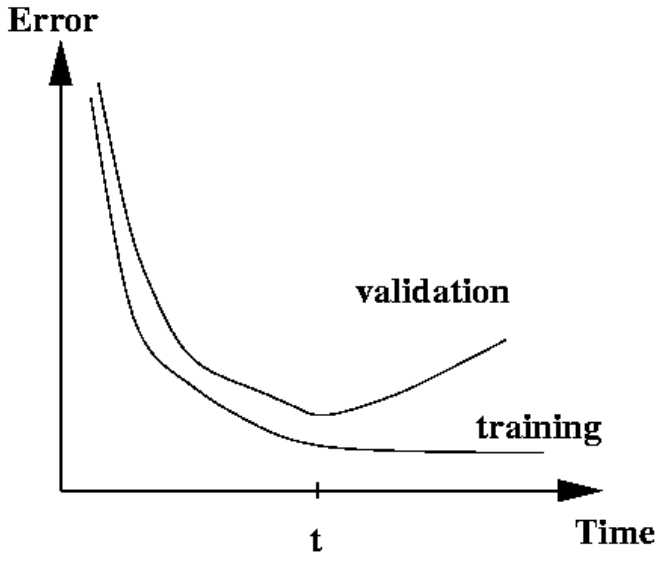


$$Accuracy = \frac{\text{Number of correct classifications}}{\text{Total number of test cases}},$$

Dataset



- **Training set**: a set of examples used for learning, where the target value is known.
- **Validation set**: a set of examples used to tune the architecture of a classifier and estimate the error.
- **Test set**: used only to assess the performances of a classifier. It is never used during the training process so that the error on the test set provides an unbiased estimate of the generalization error.



Unsupervised Machine Learning

- Given only samples X of the data, we compute a function f such that

  y = f(X) is "simpler".

  - **Clustering:** y is discrete
  - Y is continuous: **Matrix factorization, Kalman filtering, unsupervised neural networks**.

Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

- **Examples of Unsupervised:**
  - Cluster some hand-written digit data into 10 classes.
  - What are the top 20 topics in Twitter right now?
  - Find and cluster distinct accents of people at Berkeley.

Depending on the problem at hand, the unsupervised learning model can organize the data in different ways.

- **Clustering**: Without being an expert, it's possible to look at a collection of bird photos and separate them roughly by species, relying on cues like feather color, size or beak shape. That's how the most common application for unsupervised learning, clustering, works: the deep learning model looks for training data that are ***similar to each other and groups them together***.

- **Anomaly detection**: Banks detect fraudulent transactions by looking for ***unusual patterns*** in customer's purchasing behavior. For instance, if the same credit card is used in California and Denmark within the same day, that's cause for suspicion. Similarly, unsupervised learning can be used to ***flag outliers*** in a dataset.

- **Association**: Fill an online shopping cart with Hard Drive and GPU and the site just may recommend that you add SSD and network accessories to your order. This is an example of association, where ***certain features of a data sample correlate with other features***. By looking at a couple key attributes of a data point, an unsupervised learning model can predict the other attributes with which they're commonly associated.

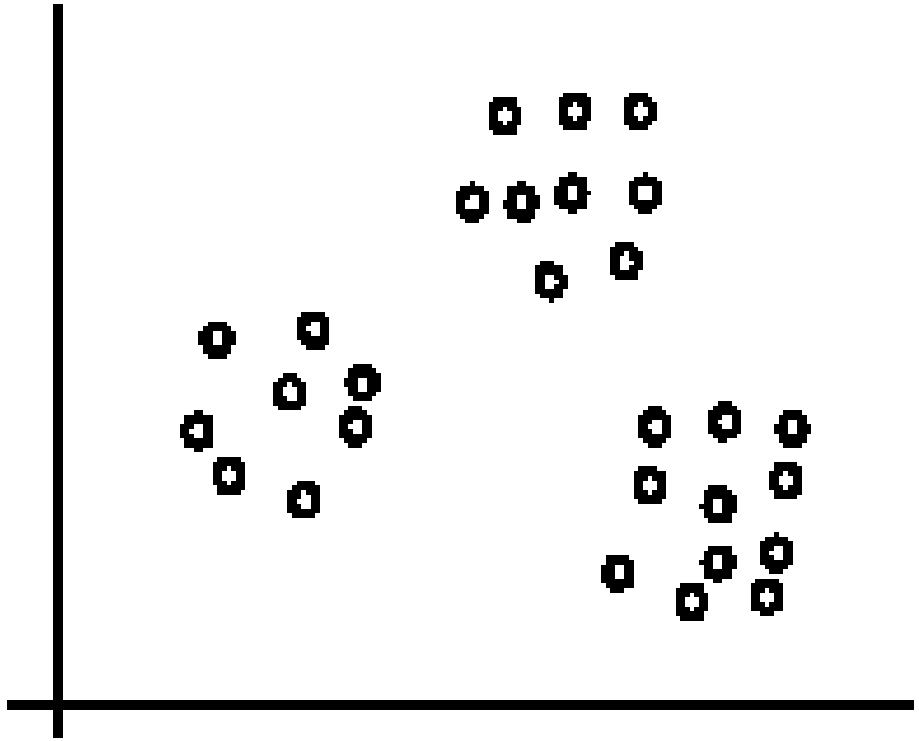
Dr. Isma F. Siddiqui *(isma.farah@faculty.muet.edu.pk)*

Clustering: Unsupervised Learning

- Clustering is a technique for finding similarity groups in data, called **clusters**. i.e.,
  - it groups data instances that are similar to (near) each other in one cluster and data instances that are very different (far away) from each other into different clusters.
- Clustering is often called an **unsupervised learning** task as no class values denoting an *a priori* grouping of the data instances are given, which is the case in supervised learning.
- Due to historical reasons, clustering is often considered synonymous with unsupervised learning.
  - In fact, clustering is one of the most utilized data mining techniques.
  - It has a long history, and used in almost every field, e.g., medicine, psychology, botany, sociology, biology, archeology, marketing, insurance, libraries, etc.
  - In recent years, due to the rapid increase of online documents, text clustering becomes important.

An illustration

- The data set has three natural groups of data points, i.e., 3 natural clusters.



Dr. Isma F. Siddiqui (isma.farah@faculty.muet.edu.pk)

What is clustering for?

- Let us see some real-life examples
- Example 1: groups people of similar sizes together to make “small”, “medium” and “large” T-Shirts.
  - Tailor-made for each person: too expensive
  - One-size-fits-all: does not fit all.
- Example 2: In marketing, segment customers according to their similarities
  - To do targeted marketing.
- Example 3: Given a collection of text documents, we want to organize them according to their content similarities,
  - To produce a topic hierarchy

Dr. Isma F. Siddiqui (isma.farah@faculty.muet.edu.pk)