Pattern Recognition

What is Pattern Recognition?

Pattern Recognition is finding the similarities or patterns among small, decomposed problems that can help us solve more complex problems more efficiently.

While that sounds complicated, we actually use pattern recognition all the time to figure out every day problems! For example, have you sorted matching pairs of socks? This is a form of pattern recognition because you used the features of the socks in order to sort them into groups.

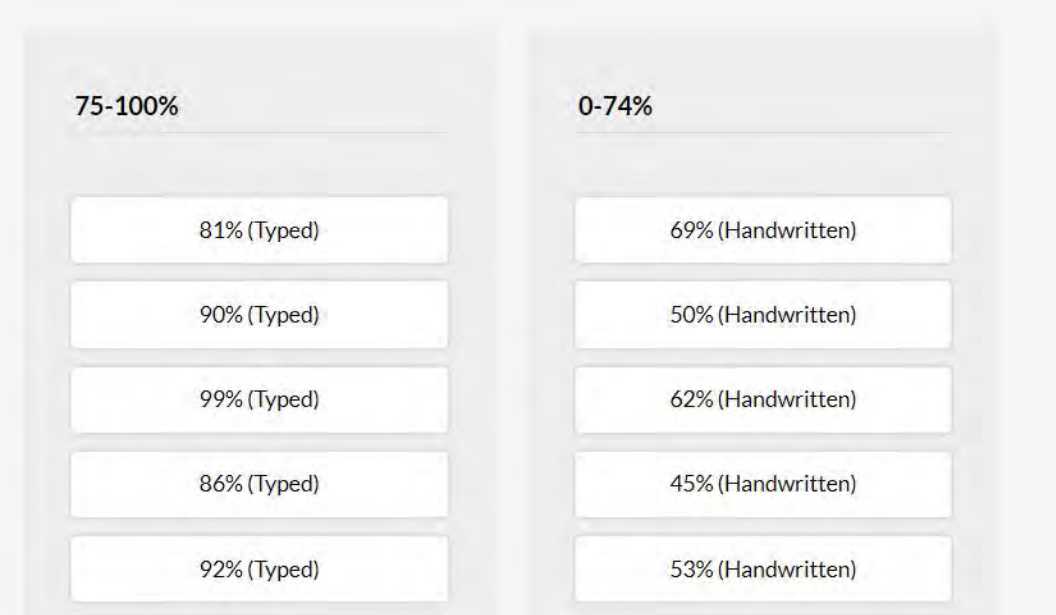
Learning Objectives

At the end of this module, you will be able to:

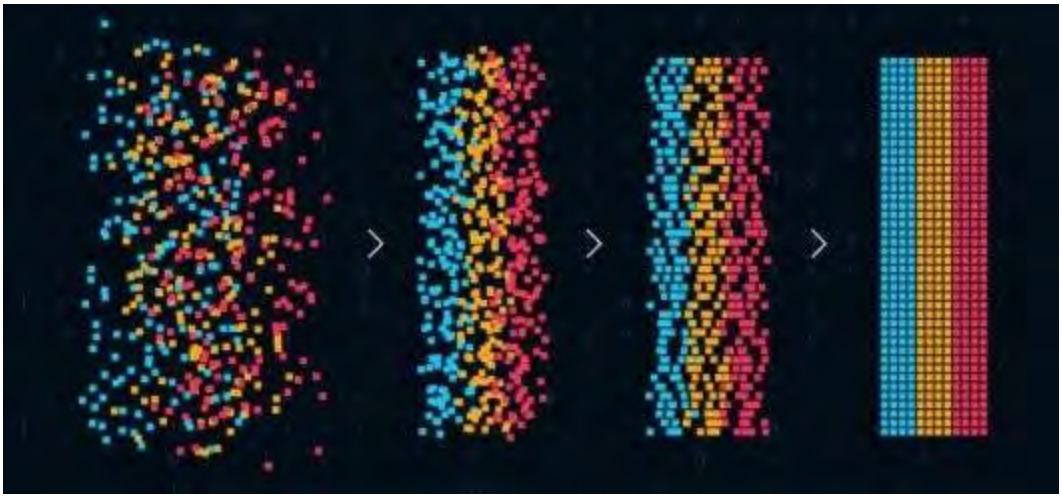
1. Define Pattern Recognition.
2. Apply Pattern Recognition to complex problems.
3. Differentiate between Pattern Recognition in Computer Science and Computer Thinking.
4. Differentiate between clustering and classification.

The first step in pattern recognition is **decomposition**

Now imagine you are a high school English teacher. You have noticed that some of your students are strong writers, while others are not. Sort the essays below to see if you notice a pattern.



Did you notice a pattern? It would seem that the students who type their papers get much better grades than their peers who decide to write their essays by hand. Now that you have noticed this pattern, you have a better idea as to your problem. Maybe some of your students went to a writing camp where typing was required? Is your grading biased? Either way, by sorting the data using pattern recognition, it is much easier to ask questions.



The Basics of Pattern Recognition

Pattern recognition is at its most basic level is all about recognizing similarities and differences, just like we did in the sorting activity above. When sorting through the student's grades, a clear set of similarities and differences emerged that showed us that students who hand write their essays score lower. In that case we were sorting based on attributes, but that's not the only thing we can look for in order to establish similarities and differences. Reference the list below to see all the things we can look for:

- Similarities in attributes, like colors and size
- Similarity in order, look for things like alternating red and green, or evens are large – odds are small
- Auditory patterns – does your car make a sound when the weather is a certain way?
- Patterns of texture

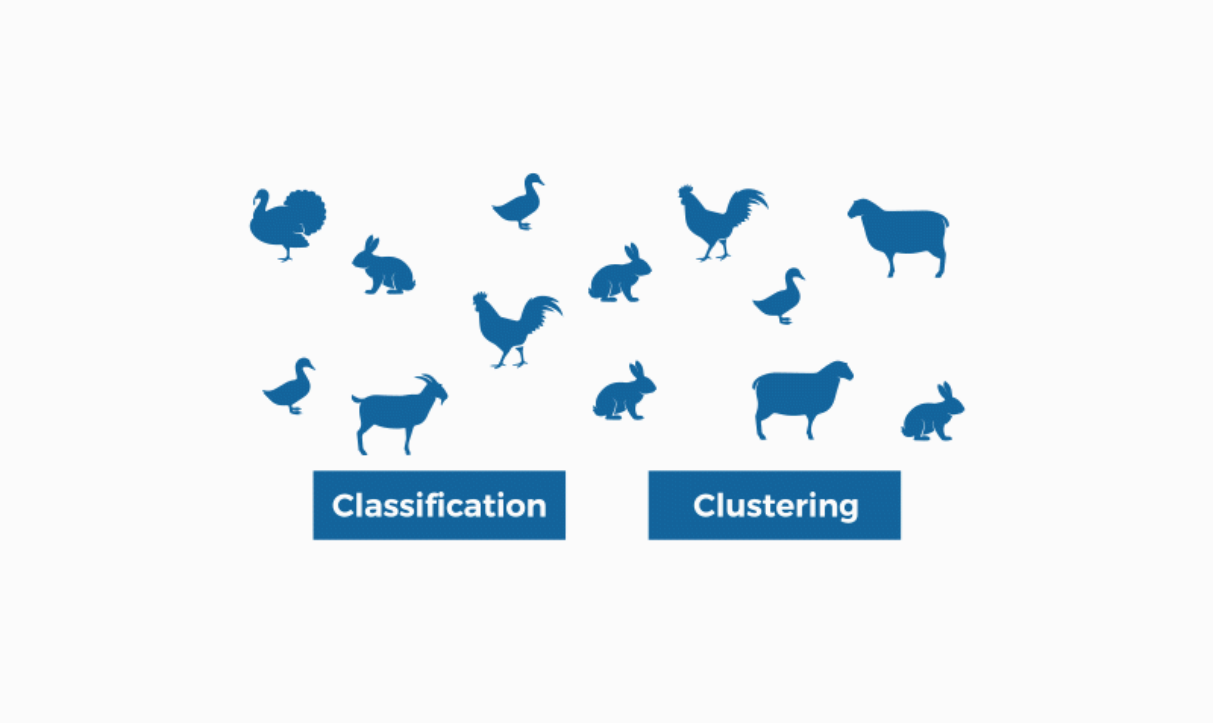
Classification and Clustering

Now that we know a pattern exists, what do we do? We need to sort it our data! There are two ways of doing this, **Classification** and **Clustering**

Classification

In classification, an appropriate class label is assigned to a pattern based on an abstraction that is generated using a set of training patterns or domain knowledge.

Essentially, classification uses predefined groups in which things are assigned into. An example of classification would be if you were told to sort all the food items into baskets labeled "vegan" and "non-vegan".



Clustering

Clustering is generating a partition of the data which helps the specific decision making activity of interest to us.

The main difference between classification and clustering is that clustering relies solely on the properties of the data provided. In classification there are instructions provided, but with clustering, this is not the case. Instead of being told to split food into vegan and non-vegan, you were told to consider all the characteristics of the food and group them together based on how similar they are.

The Steps of Pattern Recognition

1 Collect data that may be relevant. In general, this can involve collecting data corresponding to many characteristics. This task can be difficult; however, you often have

some insight into what the problem may be. For example, an experienced physician will be able to recognize commonalities.

2. Cluster the data to find commonality.

3. Determine whether those clusters are significant.

Pattern Recognition in CS and CT

While the basics of pattern recognition stay the same, it can have different applications based on the field it is being used in. For our purposes, we are going to be looking at the applications of pattern recognition in Computational Thinking and Computer Science

Computational Thinking

Specifically, with **computational thinking**, pattern recognition occurs as people study the different decomposed problems. Through analysis, students recognize patterns or connections among the different pieces of the larger problem. These patterns can be both shared similarities and shared differences. This concept is essential to building understanding amid dense information and goes well beyond recognizing patterns amongst sequences of numbers, characters, or symbols.

Computer Science

Pattern recognition in **computer science** is the imposition of identity on input data, such as speech, images, or a stream of text, by the recognition and delineation of patterns it contains and their relationships. Stages in pattern recognition may involve measurement of the object to identify

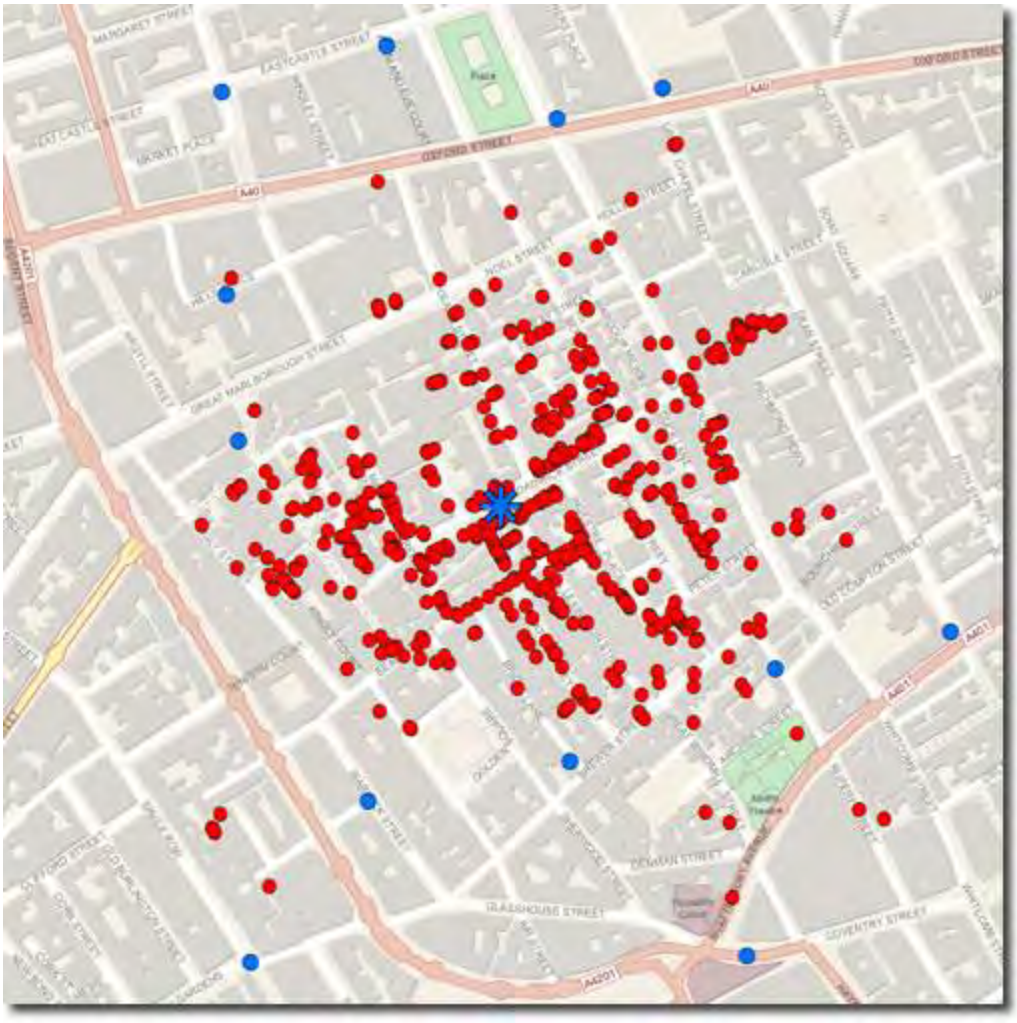
distinguishing attributes, extraction of features for the defining attributes, and comparison with known patterns to determine a match or mismatch.

Example 1

- Suppose that you were a small-animal veterinarian who noticed a large increase in disease in cats in your geographical area.
    - What characteristics might you examine to look for a pattern?
    - For example:
        - Are the diseased cats all of a similar age?
        - Are they all male? Or all female?
        - Are they all short-haired?
    - What other characteristics might you examine?
- How would you organize this data to look for clusters?
    - You might graph the data in some way

Example 2

The year is 1854 and Dr. John Snow is trying to figure out the source of a Cholera outbreak in London, England. He (and some other scientists and physicians) thought that drinking water might be the source, but not everyone was convinced of this. He mapped the homes of cholera victims and common water sources (wells) in order to prove his point. This is what they found:



The blue dots represent wells and the red dots represent the homes of cholera patients. As you might guess, the well in the center of the all the infected houses might have something to do with the cholera outbreak.

Example 3

- Suppose that you observe that the grades of one of your students have recently fallen dramatically.
    - What characteristics of this students might you examine to look for patterns?
        - E.g., are the issues with understanding of language
        - Are the issues primarily in one area, say math
    - Has the student had any health problems?
    - Is there an increase in stress at home for the student?
    - Does the student have more problems in the afternoon than in the morning, or vice versa

Example 4

Here is a cool website with a fun story about pattern recognition!

The Cube Rule

The Cube Rule of Food Identification as seen in first, we asked: are hot dogs sandwiches? then, the wars began The SandwichAlignment Chart tried to bring order but only chaos ensued then spake the holy prophet @Phosphatide blessing us with the grand unified theory of food identification identify any food purely by the location of structural starch ② Sandwich popular examples toast (made and photographed by me, Ryan.

READ MORE CUBERULE >

Knowledge Check

This knowledge check is an optional opportunity to see how much you've learned!

Question

01/04

What is pattern recognition?

- Pattern Recognition is finding the similarities or patterns among small, decomposed problems that can help us solve more complex problems more efficiently.
- Pattern Recognition is the process of creating categories to put data in to.
- Pattern recognition is when a computer creates a series of 0's and 1's in order to process information.
- Pattern recognition involves making observations about how powerful computer hardware is.

Question

02/04

Which statements are true of pattern recognition in computational thinking? (more than one may be correct)

- [ ] Pattern Recognition occurs as people study different decomposed problems.
- [ ] Pattern Recognition needs to be done on computers.
- [ ] Pattern Recognition only considers the similarities, and not the differences in data.
- [ ] Pattern Recognition in Computational Thinking goes far beyond numbers, character, and symbols.

Question

03/04

Which are true of pattern recognition in computer science? (more than one may be correct)

- [ ] In order to use pattern recognition software on a computer, you need a large amount of processing power; otherwise, the process could take months.

- [ ] Pattern recognition in computer science is the imposition of identity on input data, such as speech, images, or a stream of text, by the recognition and delineation of patterns it contains and their relationships.

- [ ] Distinguishing attributes of data are not important when doing pattern recognition in the context of computer science.

- [ ] Stages in pattern recognition may involve measurement of the object to identify distinguishing attributes, extraction of features for the defining attributes, and comparison with known patterns to determine a match or mismatch.

Question

04/04

In _____, an appropriate class label is assigned to a pattern based on an abstraction that is generated using a set of training patterns or domain knowledge.

Type your answer here