**Cognitive Computing (Perception, Learning, Reasoning)**

AI is at the forefront of a new era of computing, Cognitive Computing. It's a radically new kind of computing, very different from the programmable systems that preceded it, as different as those systems were from the tabulating machines of a century ago. Conventional computing solutions, based on the mathematical principles that emanate from the 1940's, are programmed based on rules and logic intended to derive mathematically precise answers, often following a rigid decision tree approach. But with today's wealth of big data and the need for more complex evidence-based decisions, such a rigid approach often breaks or fails to keep up with available information. Cognitive Computing enables people to create a profoundly new kind of value, finding answers and insights locked away in volumes of data. Whether we consider a doctor diagnosing a patient, a wealth manager advising a client on their retirement portfolio, or even a chef creating a new recipe, they need new approaches to put into context the volume of information they deal with on a daily basis in order to derive value from it. These processes serve to enhance human expertise. Cognitive Computing mirrors some of the key cognitive elements of human expertise, systems that reason about problems like a human does. When we as humans seek to understand something and to make a decision, we go through four key steps. First, we observe visible phenomena and bodies of evidence. Second, we draw on what we know to interpret what we are seeing to generate hypotheses about what it means. Third, we evaluate which hypotheses are right or wrong. Finally, we decide, choosing the option that seems best and acting accordingly. Just as humans become experts by going through the process of observation, evaluation, and decision-making, cognitive systems use similar processes to reason about the information they read, and they can do this at massive speed and scale. Unlike conventional computing solutions, which can only handle neatly organized structured data such as what is stored in a database, cognitive computing solutions can understand unstructured data, which is 80 percent of data today. All of the information that is produced primarily by humans for other humans to consume. This includes everything from literature, articles, research reports to blogs, posts, and tweets. While structured data is governed by well-defined fields that contain well-specified information, cognitive systems rely on natural language, which is governed by rules of grammar, context, and culture. It is implicit, ambiguous, complex, and a challenge to process. While all human language is difficult to parse, certain idioms can be particularly challenging. In English for instance, we can feel blue because it's raining cats and dogs, while we're filling in a form, someone asked us to fill out. Cognitive systems read and interpret text like a person. They do this by breaking down a sentence grammatically, relationally, and structurally, discerning meaning from the semantics of the written material. Cognitive systems understand context. This is very different from simple speech recognition, which is how a computer translates human speech into a set of words. Cognitive systems try to understand the real intent of the users language, and use that understanding to draw inferences through a broad array of linguistic models and algorithms. Cognitive systems learn, adapt, and keep getting smarter. They do this by learning from their interactions with us, and from their own successes and failures, just like humans do.

Lesson Summary

In this lesson, you have learned:

Cognitive computing systems differ from conventional computing systems in that they can:

Read and interpret unstructured data, understanding not just the meaning of words but also the intent and context in which they are used.

Reason about problems in a way that humans reason and make decisions.

Learn over time from their interactions with humans and keep getting smarter.

To learn more about how to get started with cognitive computing, read this post:

How to get started with cognitive technology

**Terminology and Related Concepts**

Before we deep dive into how AI works, and its various use cases and applications, let's differentiate some of the closely related terms and concepts of AI: artificial intelligence, machine learning, deep learning, and neural networks. These terms are sometimes used interchangeably, but they do not refer to the same thing. Artificial intelligence is a branch of computer science dealing with a simulation of intelligent behavior. AI systems will typically demonstrate behaviors associated with human intelligence such as planning, learning, reasoning, problem-solving, knowledge representation, perception, motion, and manipulation, and to a lesser extent social intelligence and creativity. Machine learning is a subset of AI that uses computer algorithms to analyze data and make intelligent decisions based on what it has learned, without being explicitly programmed. Machine learning algorithms are trained with large sets of data and they learn from examples. They do not follow rules-based algorithms. Machine learning is what enables machines to solve problems on their own and make accurate predictions using the provided data. Deep learning is a specialized subset of Machine Learning that uses layered neural networks to simulate human decision-making. Deep learning algorithms can label and categorize information and identify patterns. It is what enables AI systems to continuously learn on the job, and improve the quality and accuracy of results by determining whether decisions were correct. Artificial neural networks often referred to simply as neural networks take inspiration from biological neural networks, although they work quite a bit differently. A neural network in AI is a collection of small computing units called neurons that take incoming data and learn to make decisions over time. Neural networks are often layered deep and are the reason deep learning algorithms become more efficient as the datasets increase in volume, as opposed to other machine learning algorithms that may plateau as data increases. Now that you have a broad understanding of the differences between some key AI concepts, there is one more differentiation that is important to understand, that between artificial intelligence and data science. Data science is the process and method for extracting knowledge and insights from large volumes of disparate data. It's an interdisciplinary field involving mathematics, statistical analysis, data visualization, machine learning, and more. It's what makes it possible for us to appropriate information, see patterns, find meaning from large volumes of data, and use it to make decisions that drive business. Data Science can use many of the AI techniques to derive insight from data. For example, it could use machine learning algorithms and even deep learning models to extract meaning and draw inferences from data. There is some intersection between AI and data science, but one is not a subset of the other. Rather, data science is a broad term that encompasses the entire data processing methodology. Well, AI includes everything that allows computers to learn how to solve problems and make intelligent decisions. Both AI and Data Science can involve the use of big data that is significantly large volumes of data. In the next few lessons, the terms machine learning, deep learning, and neural networks will be discussed in more detail.

**Machine Learning**

Machine Learning, a subset of AI, uses computer algorithms to analyze data and make intelligent decisions based on what it has learned. Instead of following rules-based algorithms, machine learning builds models to classify and make predictions from data. Let's understand this by exploring a problem we may be able to tackle with Machine Learning. What if we want to determine whether a heart can fail, is this something we can solve with Machine Learning? The answer is, Yes. Let's say we are given data such as beats per minute, body mass index, age, sex, and the result whether the heart has failed or not. With Machine Learning given this dataset, we are able to learn and create a model that given inputs, will predict results. So what is the difference between this and using statistical analysis to create an algorithm? An algorithm is a mathematical technique. With traditional programming, we take data and rules, and use these to develop an algorithm that will give us an answer. In the previous example, if we were using a traditional algorithm, we would take the data such as beats per minute and BMI, and use this data to create an algorithm that will determine whether the heart will fail or not. Essentially, it would be an if-then-else statement. When we submit inputs, we get answers based on what the algorithm we determined is, and this algorithm will not change. Machine Learning, on the other hand, takes data and answers and creates the algorithm. Instead of getting answers in the end, we already have the answers. What we get is a set of rules that determine what the machine learning model will be. The model determines the rules, and the if-then-else statement when it gets the inputs. Essentially, what the model does is determine what the parameters are in a traditional algorithm, and instead of deciding arbitrarily that beats per minute plus BMI equals a certain result, we use the model to determine what the logic will be. This model, unlike a traditional algorithm, can be continuously trained and be used in the future to predict values. Machine Learning relies on defining behavioral rules by examining and comparing large datasets to find common patterns. For instance, we can provide a machine learning program with a large volume of pictures of birds and train the model to return the label "bird" whenever it has provided a picture of a bird. We can also create a label for "cat" and provide pictures of cats to train on. When the machine model is shown a picture of a cat or a bird, it will label the picture with some level of confidence. This type of Machine Learning is called Supervised Learning, where an algorithm is trained on human-labeled data. The more samples you provide a supervised learning algorithm, the more precise it becomes in classifying new data. Unsupervised Learning, another type of machine language, relies on giving the algorithm unlabeled data and letting it find patterns by itself. You provide the input but not labels, and let the machine infer qualities that algorithm ingests unlabeled data, draws inferences, and finds patterns. This type of learning can be useful for clustering data, where data is grouped according to how similar it is to its neighbors and dissimilar to everything else. Once the data is clustered, different techniques can be used to explore that data and look for patterns. For instance, you provide a machine learning algorithm with a constant stream of network traffic and let it independently learn the baseline, normal network activity, as well as the outlier and possibly malicious behavior happening on the network. The third type of machine learning algorithm, Reinforcement Learning, relies on providing a machine learning algorithm with a set of rules and constraints, and letting it learn how to achieve its goals. You define the state, the desired goal, allowed actions, and constraints. The algorithm figures out how to achieve the goal by trying different combinations of allowed actions, and is rewarded or punished depending on whether the decision was a good one. The algorithm tries its best to maximize its rewards within the constraints provided. You could use Reinforcement Learning to teach a machine to play chess or navigate an obstacle course.

**Machine Learning Techniques and Training**

Machine Learning is a broad field and we can split it up into three different categories, Supervised Learning, Unsupervised Learning, and Reinforcement Learning. There are many different tasks we can solve with these. Supervised Learning refers to when we have class labels in the dataset and we use these to build the classification model. What this means is when we receive data, it has labels that say what the data represents. In a previous example, we had a table with labels such as age or sex. With Unsupervised Learning, we don't have class labels and we must discover class labels from unstructured data. This could involve things such as deep learning looking at pictures to train models. Things like this are typically done with something called clustering. Reinforcement Learning is a different subset, and what this does is it uses a reward function to penalize bad actions or reward good actions. Breaking down Supervised Learning, we can split it up into three categories, Regression, Classification and Neural Networks. Regression models are built by looking at the relationships between features x and the result y where y is a continuous variable. Essentially, Regression estimates continuous values. Neural Networks refer to structures that imitate the structure of the human brain. Classification on the other hand, focuses on discrete values it identifies. We can assign discrete class labels y based on many input features x. In a previous example, given a set of features x, like beats per minute, body mass index, age and sex, the algorithm classifies the output y as two categories, True or False, predicting whether the heart will fail or not. In other Classification models, we can classify results into more than two categories. For example, predicting whether a recipe is for an Indian, Chinese, Japanese, or Thai dish. Some forms of classification include decision trees, support vector machines, logistic regression, and random forests. With Classification, we can extract features from the data. The features in this example would be beats per minute or age. Features are distinctive properties of input patterns that help in determining the output categories or classes of output. Each column is a feature and each row is a data point. Classification is the process of predicting the class of given data points. Our classifier uses some training data to understand how given input variables relate to that class. What exactly do we mean by training? Training refers to using a learning algorithm to determine and develop the parameters of your model. While there are many algorithms to do this, in layman's terms, if you're training a model to predict whether the heart will fail or not, that is True or False values, you will be showing the algorithm some real-life data labeled True, then showing the algorithm again, some data labeled False, and you will be repeating this process with data having True or False values, that is whether the heart actually failed or not. The algorithm modifies internal values until it has learned to tell from data that indicates heart failure that is True or not, that is False. With Machine Learning, we typically take a dataset and split it into three sets, Training, Validation and Test sets. The Training subset is the data used to train the algorithm. The Validation subset is used to validate our results and fine-tune the algorithm's parameters. The Testing data is the data the model has never seen before and used to evaluate how good our model is. We can then indicate how good the model is using terms like, accuracy, precision and recall.

Which of the following are attributes of Classification?



Classification is the process of predicting the class of given data points

**Correct**

Classification is the process of extracting features from data and classifying the results into one or more categories.

Classification is the process of predicting the class of given data points

is selected.This is correct.

Classification is the process of extracting features from data and classifying the results into one or more categories.



Forms of classification include decision trees, support vector machines, logistic regression and random forests

**Correct**

Classification is the process of extracting features from data and classifying the results into one or more categories.

Forms of classification include decision trees, support vector machines, logistic regression and random forests

is selected.This is correct.

Classification is the process of extracting features from data and classifying the results into one or more categories.



Classification models are built by looking at the relationships between features and results, where results are a continuous variable

**This should not be selected**

Classification focuses on results that have discrete values or class labels. Regression on the other hand works on estimating results that are continuous values.

Classification models are built by looking at the relationships between features and results, where results are a continuous variable

is selected.This is wrong. It should not be selected.

Classification focuses on results that have discrete values or class labels. Regression on the other hand works on estimating results that are continuous values.



Using classification models we extract features from data and classify results into multiple categories

**Correct**

Classification is the process of extracting features from data and classifying the results into one or more categories.

**Deep Learning**

[MUSIC] While Machine Learning is a subset of Artificial Intelligence, Deep Learning is a specialized subset of Machine Learning. Deep Learning layers algorithms to create a Neural Network, an artificial replication of the structure and functionality of the brain, enabling AI systems to continuously learn on the job and improve the quality and accuracy of results. This is what enables these systems to learn from unstructured data such as photos, videos, and audio files. Deep Learning, for example, enables natural language understanding capabilities of AI systems, and allows them to work out the context and intent of what is being conveyed. Deep learning algorithms do not directly map input to output. Instead, they rely on several layers of processing units. Each layer passes its output to the next layer, which processes it and passes it to the next. The many layers is why it’s called deep learning. When creating deep learning algorithms, developers and engineers configure the number of layers and the type of functions that connect the outputs of each layer to the inputs of the next. Then they train the model by providing it with lots of annotated examples. For instance, you give a deep learning algorithm thousands of images and labels that correspond to the content of each image. The algorithm will run the those examples through its layered neural network, and adjust the weights of the variables in each layer of the neural network to be able to detect the common patterns that define the images with similar labels. Deep Learning fixes one of the major problems present in older generations of learning algorithms. While the efficiency and performance of machine learning algorithms plateau as the datasets grow, deep learning algorithms continue to improve as they are fed more data. Deep Learning has proven to be very efficient at various tasks, including image captioning, voice recognition and transcription, facial recognition, medical imaging, and language translation. Deep Learning is also one of the main components of driverless cars. [MUSIC]

**Neural Networks**

An artificial neural network is a collection of smaller units called neurons, which are computing units modeled on the way the human brain processes information. Artificial neural networks borrow some ideas from the biological neural network of the brain, in order to approximate some of its processing results. These units or neurons take incoming data like the biological neural networks and learn to make decisions over time. Neural networks learn through a process called backpropagation. Backpropagation uses a set of training data that match known inputs to desired outputs. First, the inputs are plugged into the network and outputs are determined. Then, an error function determines how far the given output is from the desired output. Finally, adjustments are made in order to reduce errors. A collection of neurons is called a layer, and a layer takes in an input and provides an output. Any neural network will have one input layer and one output layer. It will also have one or more hidden layers which simulate the types of activity that goes on in the human brain. Hidden layers take in a set of weighted inputs and produce an output through an activation function. A neural network having more than one hidden layer is referred to as a deep neural network. Perceptrons are the simplest and oldest types of neural networks. They are single-layered neural networks consisting of input nodes connected directly to an output node. Input layers forward the input values to the next layer, by means of multiplying by a weight and summing the results. Hidden layers receive input from other nodes and forward their output to other nodes. Hidden and output nodes have a property called bias, which is a special type of weight that applies to a node after the other inputs are considered. Finally, an activation function determines how a node responds to its inputs. The function is run against the sum of the inputs and bias, and then the result is forwarded as an output. Activation functions can take different forms, and choosing them is a critical component to the success of a neural network. Convolutional neural networks or CNNs are multilayer neural networks that take inspiration from the animal visual cortex. CNNs are useful in applications such as image processing, video recognition, and natural language processing. A convolution is a mathematical operation, where a function is applied to another function and the result is a mixture of the two functions. Convolutions are good at detecting simple structures in an image, and putting those simple features together to construct more complex features. In a convolutional network, this process occurs over a series of layers, each of which conducts a convolution on the output of the previous layer. CNNs are adept at building complex features from less complex ones. Recurrent neural networks or RNNs, are recurrent because they perform the same task for every element of a sequence, with prior outputs feeding subsequent stage inputs. In a general neural network, an input is processed through a number of layers and an output is produced with an assumption that the two successive inputs are independent of each other, but that may not hold true in certain scenarios. For example, when we need to consider the context in which a word has been spoken, in such scenarios, dependence on previous observations has to be considered to produce the output. RNNs can make use of information in long sequences, each layer of the network representing the observation at a certain time.

Paint with AI

IBM Research creates innovative tools and resources to help unleash the power of AI. In this hands on lab, you will use a new kind of neural network, called a generative adversarial network (GAN), to create complex outputs, like photorealistic images. You will use a GAN to enhance existing images and create your own unique, custom image.

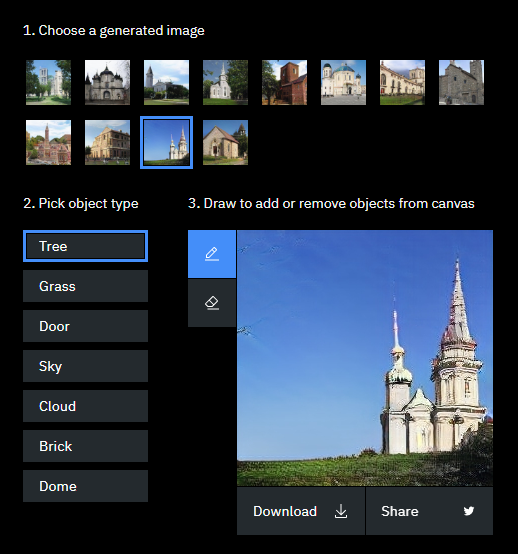
Follow these steps to work with a GAN:

1. Access the demo here: This landscape image was created by AI.

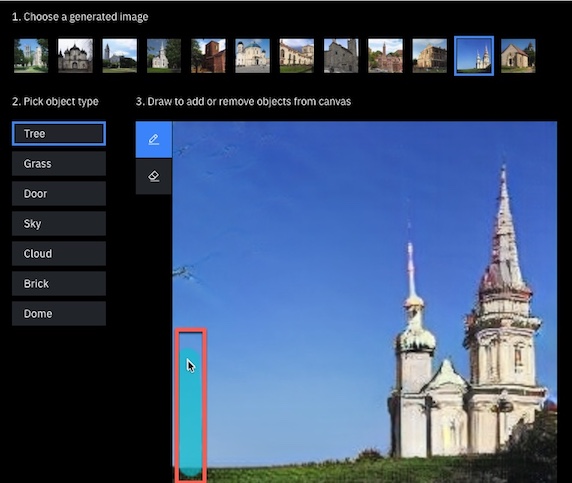
2. In the Co-create with a neural network section, under Choose a generated image, select one of the existing images. For example, choose the 11th image.

3. From the Pick object type list, select the type of object you want to add. For example, click on Tree to select it.

Figure 1 - Original generated image:

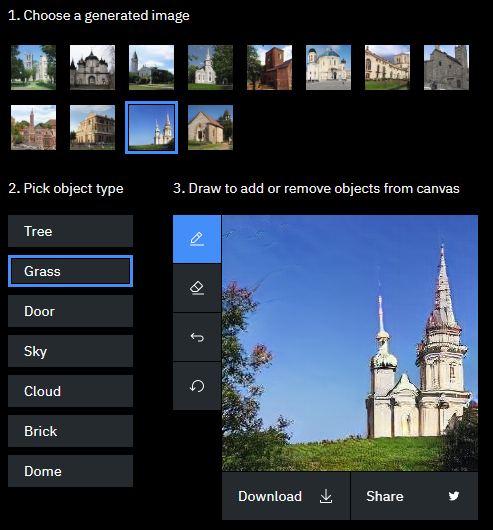


4. Move the cursor onto the image. Click and keeping the mouse button pressed, drag your cursor over an area of the existing image where you want to add the object. For example drag a line in the red area highlighted in the red rectangle below to add a tree there.



5. Choose another object type and add it to the image.

Figure 2 - Trees and grass added:



6. Experiment with locations: Can you place a door in the sky? Can you place grass so that it enhances the appearance of existing grass? 7. Use the Undo and Erase functions to remove objects. 8. [Optional] Click Download to save your work.

For more information on the capabilities of GANs, follow these steps:

In the What’s happening in this demo section? Click What does a GAN “understand” and read the text.

What does the text say about placement of objects? Does this explain the results you saw earlier?

Click Painting with neurons, not pixels and read the text. How does the GAN help you manipulate images?

Click New ways to work with AI and read the text. What are some of the use cases for GANs?

Use the Discussion Forum to talk about these questions with your fellow students.

Lesson Summary

In this lesson, you have learned:

* Machine Learning, a subset of AI, uses computer algorithms to analyze data and make intelligent decisions based on what it has learned. The three main categories of machine learning algorithms include Supervised Learning, Unsupervised Learning, and Reinforcement learning.
* Deep Learning, a specialized subset of Machine Learning, layers algorithms to create a neural network enabling AI systems to learn from unstructured data and continue learning on the job.
* Neural Networks, a collection of computing units modeled on biological neurons, take incoming data and learn to make decisions over time. The different types of neural networks include Perceptrons, Convolutional Neural Networks or CNNs, and Recurrent Neural Networks or RNNs.

In the Machine Learning Techniques and Training topic, you have learned:

* Supervised Learning is when we have class labels in the data set and use these to build the classification model.
* Supervised Learning is split into three categories – Regression, Classification, and Neural Networks.
* Machine learning algorithms are trained using data sets split into training data, validation data, and test data.

[Optional] To learn more about Machine Learning, Deep Learning, and Neural Networks, read these articles:

* [Models for Machine Learning](https://developer.ibm.com/articles/cc-models-machine-learning/)
* [Applications of Deep Learning](https://www.ibm.com/blogs/think/category/deep-learning)
* [A Neural Networks deep dive](https://developer.ibm.com/articles/cc-cognitive-neural-networks-deep-dive/?mhq=neural%20networks&mhsrc=ibmsearch_a)

**Key Fields of Application in AI**

So can you talk about the different areas or categories of artificial intelligence? Now, there are lots of different fields that AI works in. But if I were to on a very very high level group some of the major areas where artificial intelligence is applied, I'd like to start off with natural language. Because natural language is, I'd say, the most complex data for machine learning to work with. If you see all sorts of data, whether that be a sequence to genome, whether that be audio, whether that be images. There's some sort of discernible pattern. There's some sort of yes, this is what a car sounds like or yes, this is what human voice sounds like. But natural language is fundamentally, a very human task. It's very human data source. We as humans invented it for humans to understand. If I were to, for example, give you a book title, there's actually a very very famous book, and the title of the book is there are two mistakes in the the title of this book. Now, there's actually only one mistake, the two the's. The human brain doesn't realize that. What's the second mistake? That there was only one mistake. So this is a sort of natural language complexity that's involved here. Humans we don't view natural language literally. We view it conceptually. If I were to write a three instead of an E, you will understand it because we don't mean the three in a literal sense. We mean that in a symbolic sense to represent the concept of E and you can contextualize that three to figure out that, "Yeah. It means in E" and not an actual three. These are things that computers aren't capable of. So natural languages that number one field that I'm most interested in when it comes to machine learning. Second, I'd say the most popular would be visual. Visual data understanding, computer vision. Because it enables us to do so many things. As humans, our primary sense is vision. In fact, a vast majority of your brain's processing power at any given moment, goes to understanding what it is that you're seeing. Whether it be a person's face, or whether it be a computer or some texts, or anything of that sort. Third, I would say audio-based data. So text-to-speech, speech-to-text these are very very complex. The reason it's complex is because it combines a lot of challenges into one. First of all, you've got to support many languages. You can't just support English and call it a day. You've got to support other languages. You've got to support other demographics. Another challenge is that even within languages, there are absolutely infinite number of ways that any human could represent a language. Everyone's going to have a different accent. Everyone's going to have a different way of pronouncing certain words. There's no standardized way that every human will pronounce ice cube exactly like ice cube. That doesn't exist. If you take a look at another challenge, it's that audio data is fundamentally very very difficult to work with. Because the thing is, audio data exists in the natural world. What is audio? It's vibrations of air molecules, and vibrations of air molecules are fast. Audio is recorded at overpay say 44 kilohertz. That's a lot of data, 44,000 data points every single second. There are usually only 44,000 data points in an individual low-resolution image. So of course, there are lots of challenges to work around when it comes to audio. But companies like IBM, Google, Microsoft have actually worked around these challenges and they're working towards creating different services to make it easier for developers. So again, on a very very high level, there's natural language understanding, there's computer vision, there's audio data and of course, there's the traditional set of tabular data understanding. Which is essentially, structured data understanding.

**Natural Language Processing, Speech, Computer Vision**

Some of the most common application areas of AI include natural language processing, speech, and computer vision. Now, let's look at each of these in turn. Humans have the most advanced method of communication which is known as natural language. While humans can use computers to send voice and text messages to each other, computers do not innately know how to process natural language. Natural language processing is a subset of artificial intelligence that enables computers to understand the meaning of human language. Natural language processing uses machine learning and deep learning algorithms to discern a word's semantic meaning. It does this by deconstructing sentences grammatically, relationally, and structurally and understanding the context of use. For instance, based on the context of a conversation, NLP can determine if the word "Cloud" is a reference to cloud computing or the mass of condensed water vapor floating in the sky. NLP systems might also be able to understand intent and emotion, such as whether you're asking a question out of frustration, confusion, or irritation. Understanding the real intent of the user's language, NLP systems draw inferences through a broad array of linguistic models and algorithms. Natural language processing is broken down into many subcategories related to audio and visual tasks. For computers to communicate in natural language, they need to be able to convert speech into text, so communication is more natural and easy to process. They also need to be able to convert text-to-speech, so users can interact with computers without the requirement to stare at a screen. The older iterations of speech-to-text technology require programmers to go through tedious process of discovering and codifying the rules of classifying and converting voice samples into text. With neural networks, instead of coding the rules, you provide voice samples and their corresponding text. The neural network finds the common patterns among the pronunciation of words and then learns to map new voice recordings to their corresponding texts. These advances in speech-to-text technology are the reason we have real time transcription. Google uses AI-powered speech-to-text in there Call Screen feature to handle scam calls and show you the text of the person speaking in real time. YouTube uses this to provide automatic closed captioning. The flip side of speech-to-text is text-to-speech also known as speech synthesis. In the past, the creation of a voice model required hundreds of hours of coding. Now, with the help of neural networks, synthesizing human voice has become possible. First, a neural network ingests numerous samples of a person's voice until it can tell whether a new voice sample belongs to the same person. Then, a second neural network generates audio data and runs it through the first network to see if it validates it as belonging to the subject. If it does not, the generator corrects its sample and reruns it through the classifier. The two networks repeat the process until they generate samples that sound natural. Companies use AI-powered voice synthesis to enhance customer experience and give their brands their unique voice. In the medical field, this technology is helping ALS patients regain their true voice instead of using a computerized voice. The field of computer vision focuses on replicating parts of the complexity of the human visual system, and enabling computers to identify and process objects in images and videos, in the same way humans do. Computer vision is one of the technologies that enables the digital world to interact with the physical world. The field of computer vision has taken great leaps in recent years and surpasses humans in tasks related to detecting and labeling objects, thanks to advances in deep learning and neural networks. This technology enables self-driving cars to make sense of their surroundings. It plays a vital role in facial recognition applications allowing computers to match images of people's faces to their identities. It also plays a crucial role in augmented and mixed reality. The technology that allows computing devices such as smartphones, tablets, and smart glasses to overlay and embed virtual objects on real-world imagery. Online photo libraries like Google Photos, use computer vision to detect objects and classify images by the type of content they contain.

Which of the following is an attribute of Natural Language Processing (NLP)?



NLP systems are provided recorded voice samples with corresponding text to help them discern common patterns

**This should not be selected**

Speech-to-text technologies use recorded voice samples and corresponding text to find common patterns among the pronunciation of words. The model then 'learns' to map new voice recordings to their corresponding texts. NLP enables machines to understand the meaning of human language.

NLP systems are provided recorded voice samples with corresponding text to help them discern common patterns

is selected.This is wrong. It should not be selected.

Speech-to-text technologies use recorded voice samples and corresponding text to find common patterns among the pronunciation of words. The model then 'learns' to map new voice recordings to their corresponding texts. NLP enables machines to understand the meaning of human language.



NLP systems can understand intent

**Correct**

Natural Language Processing uses machine learning and deep learning algorithms to deconstruct sentences grammatically, relationally, and structurally, to understand a word's semantic meaning and it's context of use. NLP systems are also able to understand intent and emotion, by drawing inferences through a broad array of linguistic models and algorithms.

NLP systems can understand intent

is selected.This is correct.

Natural Language Processing uses machine learning and deep learning algorithms to deconstruct sentences grammatically, relationally, and structurally, to understand a word's semantic meaning and it's context of use. NLP systems are also able to understand intent and emotion, by drawing inferences through a broad array of linguistic models and algorithms.



NLP systems can identify the emotion in which a word is spoken, for example, frustration, confusion, irritation, or fun etc.

**Correct**

Natural Language Processing uses machine learning and deep learning algorithms to deconstruct sentences grammatically, relationally, and structurally, to understand a word's semantic meaning and it's context of use. NLP systems are also able to understand intent and emotion, by drawing inferences through a broad array of linguistic models and algorithms.

NLP systems can identify the emotion in which a word is spoken, for example, frustration, confusion, irritation, or fun etc.

is selected.This is correct.

Natural Language Processing uses machine learning and deep learning algorithms to deconstruct sentences grammatically, relationally, and structurally, to understand a word's semantic meaning and it's context of use. NLP systems are also able to understand intent and emotion, by drawing inferences through a broad array of linguistic models and algorithms.



NLP systems use a broad array of linguistic models and algorithms to draw inferences from language

**Correct**

Natural Language Processing uses machine learning and deep learning algorithms to deconstruct sentences grammatically, relationally, and structurally, to understand a word's semantic meaning and it's context of use. NLP systems are also able to understand intent and emotion, by drawing inferences through a broad array of linguistic models and algorithms.

NLP systems use a broad array of linguistic models and algorithms to draw inferences from language

is selected.This is correct.

Natural Language Processing uses machine learning and deep learning algorithms to deconstruct sentences grammatically, relationally, and structurally, to understand a word's semantic meaning and it's context of use. NLP systems are also able to understand intent and emotion, by drawing inferences through a broad array of linguistic models and algorithms.

**Application of AI :Self Driving Cars**

Can you tell us a little bit about the work you're doing with self-driving cars. >> I've been working on self-driving cars for the last few years. It's a domain that's exploded, obviously, in interest since early competitions back in the 2005 domain. And what we've been working on really is putting together our own self-driving vehicle that was able to drive on public roads in the regional Waterloo last August. With the self-driving cars area, one of our key research domains is in 3D object detection. So this remains a challenging task for algorithms to perform automatically. Trying to identify every vehicle, every pedestrian, every sign that's in a driving environment. So that the vehicle can make the correct decisions about how it should move and interact with those vehicles. And so we work extensively on how we take in laser data and vision data and radar data. And then fuse that into a complete view of the world around the vehicle. >> When we think of computer vision, we usually think immediately of self-driving cars, and why is that? Well, it's because it's hard to pay attention when driving on the road, right? You can't both be looking at your smartphone and also be looking at the road at the same time. Of course, it's sometimes hard to predict what people are going to be doing on the street, as well. When they're crossing the street with their bike or skateboard, or whatnot. So it's great when we have some sort of camera or sensor that can help us detect these things and prevent accidents before they could potentially occur. And that's one of the limitations of human vision, is attention, is visual attention. So I could be looking at you, Rav, but behind you could be this delicious slice of pizza. But I can only pay attention to one or just some limited number of things at a time. But I can't attend to everything in my visual field all at once at the same time like a camera could. Or like how computer vision could potentially do so. And so that's one of the great things that cameras and computer vision is good for. Helping us pay attention to the whole world around us without having us to look around and make sure that we're paying attention to everything. And that's just in self-driving cars, so I think we all kind of have a good sense of how AI and computer vision shapes the driving and transportation industry. >> Well, self-driving cars are certainly the future. And there's tremendous interest right now in self-driving vehicles. In part because of their potential to really change the way our society works and operates. I'm very excited about being able to get into a self-driving car and read or sit on the phone on the way to work. Instead of having to pilot through Toronto traffic.

Play video starting at 2 minutes 58 seconds and follow transcript2:58

So I think they represent a really exciting step forward, but there's still lots to do. We still have lots of interesting challenges to solve in the self-driving space. Before we have really robust and safe cars that are able to drive themselves 100% of the time autonomously on our roads.

Play video starting at 3 minutes 18 seconds and follow transcript3:18

>> We've just launched our own self-driving car specialization on Coursera. And we'd be really happy to see students in this specialization also come and learn more about self-driving. It's a wonderful starting point, it gives you a really nice perspective on the different components of the self-driving software stack and how it actually works. So everywhere from how it perceives the environment, how it makes decisions and plans its way through that environment. To how it controls the vehicle and makes sure it executes those plans safely. So you'll get a nice broad sweep of all of those things from that specialization. And from there you then want to become really good and really deep in one particular area, if you want to work in this domain. Because again, there's so many layers behind this. There's so much foundational knowledge you need to start contributing that you can't go wrong. If you find something interesting, just go after it. And I am sure there'll be companies that'll need you for this. [MUSIC]

In order for a self-driving vehicle to navigate accurately, it needs to piece together a complete view of its driving environment, which it does with the help of:



Laser data

**Correct**

Self-driving vehicles fuse laser data, vision data, and radar data to create a three-dimensional view of their driving environment helping them make accurate decisions on the road.

Laser data

is selected.This is correct.

Self-driving vehicles fuse laser data, vision data, and radar data to create a three-dimensional view of their driving environment helping them make accurate decisions on the road.



Vision data

**Correct**

Self-driving vehicles fuse laser data, vision data, and radar data to create a three-dimensional view of their driving environment helping them make accurate decisions on the road.

Vision data

is selected.This is correct.

Self-driving vehicles fuse laser data, vision data, and radar data to create a three-dimensional view of their driving environment helping them make accurate decisions on the road.



Radar data

**Correct**

Self-driving vehicles fuse laser data, vision data, and radar data to create a three-dimensional view of their driving environment helping them make accurate decisions on the road.

Radar data

is selected.This is correct.

Self-driving vehicles fuse laser data, vision data, and radar data to create a three-dimensional view of their driving environment helping them make accurate decisions on the road.



Data Science

**Un-selected is correct**

Which is the biggest limitation of human vision that computer vision can help make up for?



Visual Attention

**Correct**

Visual attention is a limitation of human vision that computer vision can help makeup for. Human vision cannot attend to everything in its visual field, all at the same time, like computer vision can.

Computer Vision

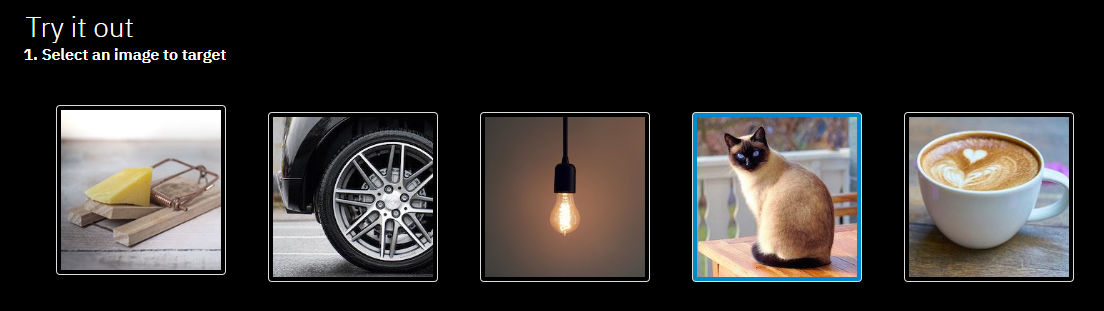
IBM Research creates innovative tools and resources to help unleash the power of AI. In this hands on lab, you will learn about IBM’s Adversarial Robustness Toolbox, and use it to mitigate simulated attacks by hackers.

Follow these steps to explore the demo:

1. Access the demo here: [Your AI model might be telling you this is not a cat.](https://art-demo.mybluemix.net/?cm_mc_uid=59963159489715526067556&cm_mc_sid_50200000=57143911561485594784&cm_mc_sid_52640000=48611511561485594790)

2. In the **Try it out** section, click the image of the Siamese cat.

*Figure 1 - Select an image*

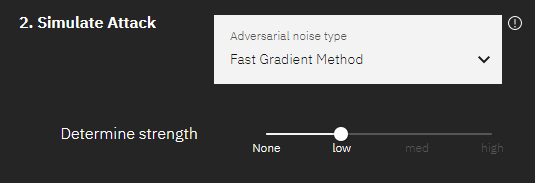


3. In the **Simulate Attack** section, ensure that no attack is selected, and that all the sliders are to the far left, indicating that all attacks and mitigation strategies are turned off.

What does Watson identify the image as, and at what confidence level? E.g. Siamese cat 92%

4. In the **Simulate Attack** section, under **Adversarial noise** **type**, select **Fast Gradient Method**. The strength slider will move to low.

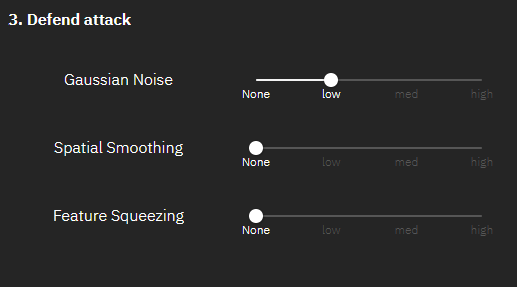
*Figure 2 - Select an attack and level*



What does Watson identify the image as now, and at what confidence level?

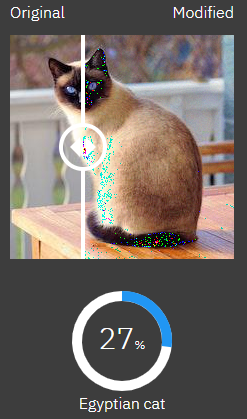
3. In the **Defend attack** section, move the **Gaussian Noise**slider to low.

*Figure 3 - Mitigate the attack*



4. What does Watson identify the image as now, and at what confidence level? Did the image recognition improve?

*Figure 4 - View the results*



Note that you can use the slider on the image to see the original and modified images.

5. Move the **Gaussian Noise** slider to **medium**, and then to **high**. For each level, note what Watson identifies the image as, and at what confidence level. Did the image recognition improve?

6. Move the **Gaussian Noise** slider to **None**.

7. In the **Defend attack** section, move the **Spatial Smoothing** slider to **low**. What does Watson identify the image as now, and at what confidence level? Did the image recognition improve?

8. Move the **Spatial Smoothing** slider to **medium**, and then to **high**. For each level, note what Watson identifies the image as, and at what confidence level. Did the image recognition improve?

9. Move the **Spatial Smoothing** slider to **None**.

10. In the **Defend attack** section, move the **Feature Squeezing**slider to **low**. What does Watson identify the image as now, and at what confidence level? Did the image recognition improve?

11. Move the **Feature Squeezing** slider to **medium**, and then to **high**. For each level, note what Watson identifies the image as, and at what confidence level. Did the image recognition improve?

12. Which of the three defenses would you use to defend against a Fast Gradient Attack?

**Optional:**

If you have time, use the same techniques to explore the other methods of attack (Projected Gradient Descent and C&W Attack) and evaluate which method of defense works best for each. If you want, try a different image.

Use the Discussion Forum to talk about the attacks and mitigation strategies with your fellow students.

Lesson Summary

In this lesson, you have learned:

* Natural Language Processing (NLP) is a subset of artificial intelligence that enables computers to understand the meaning of human language, including the intent and context of use.
* Speech-to-text enables machines to convert speech to text by identifying common patterns in the different pronunciations of a word, mapping new voice samples to corresponding words.
* Speech Synthesis enables machines to create natural sounding voice models, including the voice of particular individuals.
* Computer Vision enables machines to identify and differentiate objects in images the same way humans do.
* Self-driving cars is an application of AI that can utilize NLP, speech, and most importantly, computer vision.

To learn more about Natural Language Processing, read this article:

* [A beginner's guide to Natural Language Processing](https://developer.ibm.com/articles/a-beginners-guide-to-natural-language-processing/)