

Voici une définition de ce qu'est l'apprentissage automatique attribuée à Arthur Samuel. Elle a défini l'apprentissage automatique comme le domaine d'étude qui permet aux ordinateurs d'apprendre sans être programmés explicitement. Samuel est célèbre parce que, dans les années 1950, il a écrit un programme de jeu de dames. Ce qui est étonnant à propos de ce programme, c'est qu'Arthur Samuel lui-même n'était pas un très bon joueur de dames. Ce qu'elle a fait, c'est qu'elle avait programmé l'ordinateur pour qu'il joue à des dizaines de jeux contre lui-même.

## Apprentissage supervisé partie 1

L'apprentissage automatique supervisé, ou plus communément, l'apprentissage supervisé, fait référence à des algorithmes qui apprennent des mappages  $x \rightarrow y$  ou des mappages d'entrée à sortie. La principale caractéristique de l'apprentissage supervisé est que vous donnez des exemples d'algorithmes d'apprentissage à partir desquels vous pouvez tirer des leçons. Cela inclut les bonnes réponses, la bonne réponse étant, je veux dire, la bonne étiquette  $y$  pour une entrée  $x$  donnée, et c'est en voyant les bonnes paires d'entrée  $x$  et de l'étiquette de sortie souhaitée  $y$  que l'algorithme d'apprentissage apprend finalement à ne prendre que l'entrée seule sans l'étiquette de sortie et donne une prédition ou une estimation raisonnablement précise de la sortie.

Input (X)	Output (Y)	Application
email	spam? (0/1)	spam filtering
audio	text transcripts	speech recognition
English	Spanish	machine translation
ad, user info	click? (0/1)	online advertising
image, radar info	position of other cars	self-driving car
image of phone	defect? (0/1)	visual inspection

Regardons quelques exemples. Si l'entrée  $x$  est un e-mail et que la sortie  $y$  est cet e-mail, spam ou non, cela vous donne votre filtre anti-spam. Ou si l'entrée est un clip audio et que la tâche de l'algorithme consiste à générer la transcription du texte, il s'agit alors d'une reconnaissance vocale. Ou si vous souhaitez saisir l'anglais et le faire sortir dans la traduction correspondante en espagnol, en arabe, en hindi, en chinois, en japonais ou autre, alors c'est de la traduction automatique.

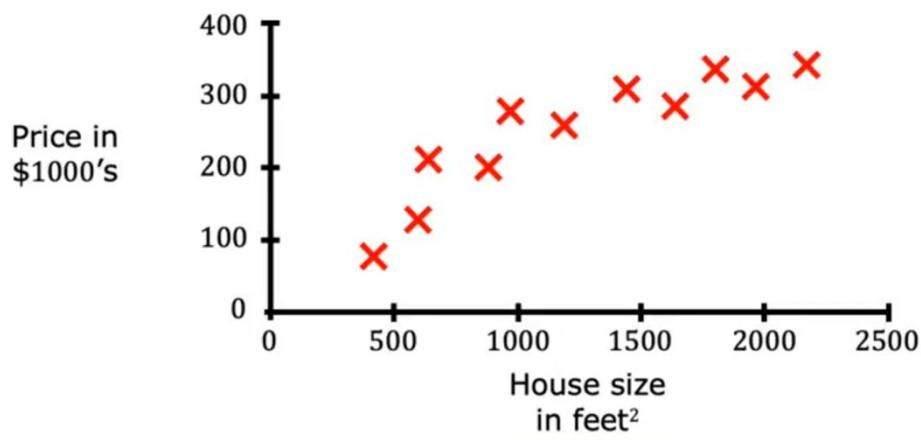
Ou encore, la forme d'apprentissage supervisé la plus lucrative aujourd'hui est probablement celle utilisée dans la publicité en ligne. Presque toutes les grandes plateformes publicitaires en ligne disposent d'un algorithme d'apprentissage qui saisit certaines informations sur une annonce et d'autres informations vous concernant, puis essaie de déterminer si vous allez cliquer sur cette annonce ou non. Parce qu'en vous diffusant des publicités sur lesquelles ils sont légèrement plus susceptibles de cliquer, pour ces grandes plateformes publicitaires en ligne, chaque clic génère des revenus, ce qui génère en fait des revenus importants pour ces entreprises.

C'est une application sur laquelle j'ai beaucoup travaillé par le passé. Ce n'est peut-être pas l'application la plus inspirante, mais elle a certainement un impact économique significatif dans certains pays aujourd'hui. Ou si vous souhaitez construire une voiture autonome, l'algorithme d'apprentissage prendra en entrée une image et certaines informations provenant d'autres capteurs, tels qu'un radar ou d'autres éléments, puis essaiera de générer la position, par exemple, d'autres voitures afin que votre voiture autonome puisse contourner les autres voitures en toute sécurité. Ou prenez l'exemple de la fabrication.

En fait, j'ai beaucoup travaillé dans ce secteur en matière d'apprentissage de l'IA. Vous pouvez demander à un algorithme d'apprentissage de prendre en entrée une image d'un produit fabriqué, par exemple un téléphone portable qui vient de sortir de la chaîne de production, et de faire sortir l'algorithme d'apprentissage, qu'il présente ou non une rayure, une bosse ou un autre défaut sur le produit. C'est ce que l'on appelle l'inspection visuelle et cela aide les fabricants à réduire ou à prévenir les défauts de leurs produits.

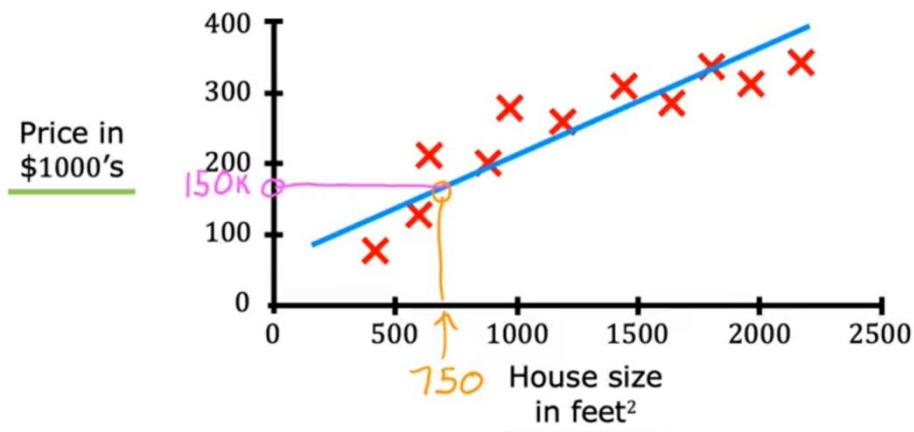
Dans toutes ces applications, vous allez d'abord entraîner votre modèle à l'aide d'exemples d'entrées  $x$  et des bonnes réponses, à savoir les étiquettes  $y$ . Une fois que le modèle a appris à partir de ces paires d'entrée, de sortie ou de ces paires  $x$  et  $y$ , il peut prendre une toute nouvelle entrée  $x$ , qu'il n'a jamais vue auparavant, et essayer de produire la sortie  $y$  correspondante appropriée.

## Regression: Housing price prediction



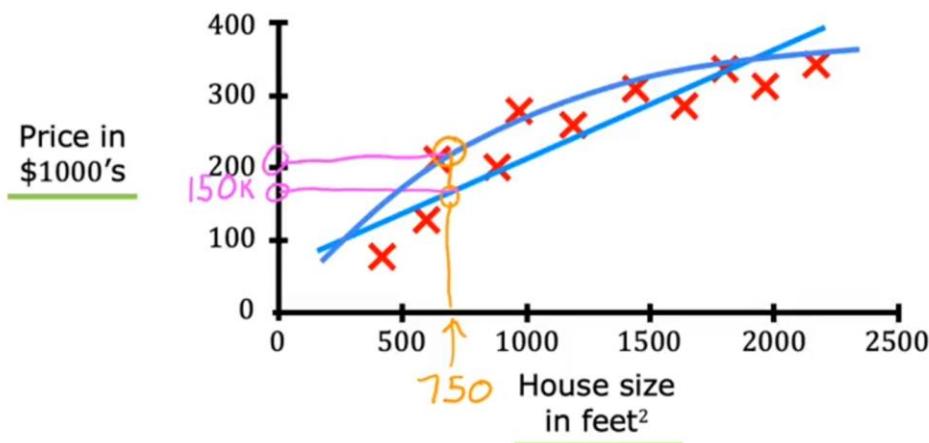
Examinons plus en détail un exemple spécifique. Supposons que vous souhaitez prévoir les prix des logements en fonction de la taille de la maison. Vous avez collecté certaines données et vous dites que vous les avez tracées, et cela ressemble à ceci. Ici, sur l'axe horizontal, se trouve la taille de la maison en pieds carrés. Oui, je vis aux États-Unis où nous utilisons toujours des pieds carrés. Je sais que la plupart des pays du monde utilisent des mètres carrés. Voici, sur l'axe vertical, le prix de la maison en milliers de dollars, par exemple. Avec ces données, supposons qu'un ami veuille savoir quel est le prix de sa maison de 750 pieds carrés. Comment l'algorithme d'apprentissage peut-il vous aider ?

## Regression: Housing price prediction



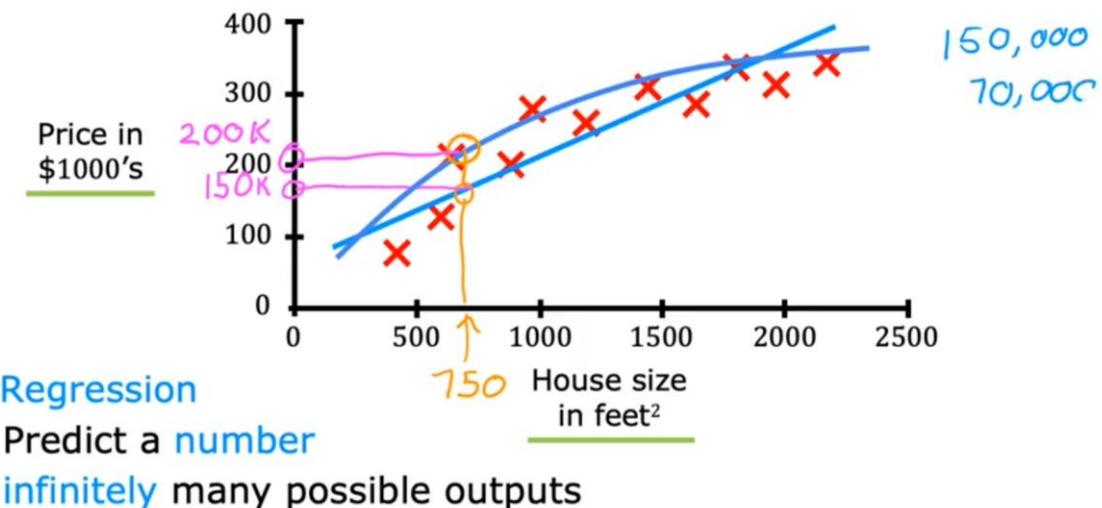
Une chose qu'un algorithme d'apprentissage pourrait être capable de faire est de dire que, pour la ligne droite menant aux données et la lecture de la ligne droite, il semblerait que la maison de votre ami pourrait être vendue pour environ, je ne sais pas, 150 000 dollars.

## Regression: Housing price prediction



Mais l'ajustement d'une ligne droite n'est pas le seul algorithme d'apprentissage que vous pouvez utiliser. Il y en a d'autres qui pourraient mieux fonctionner pour cette application. Par exemple, si vous tracez et adaptez une ligne droite, vous pouvez décider qu'il est préférable d'ajuster une courbe, une fonction légèrement plus compliquée ou plus complexe qu'une ligne droite. Si vous faites cela et que vous faites une prédiction, il semblerait que la maison de votre ami pourrait être vendue pour près de 200 000\$.

## Regression: Housing price prediction

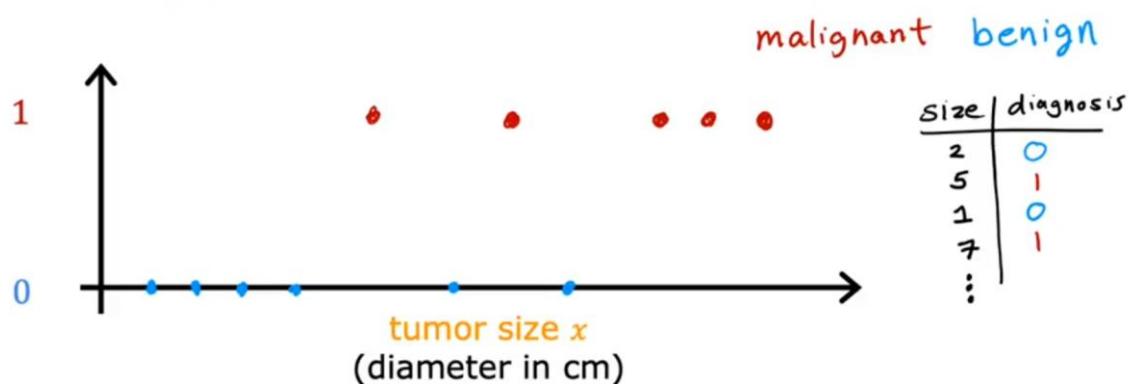


Ce que vous avez vu dans cette diapositive est un exemple d'apprentissage supervisé. Parce que nous avons fourni à l'algorithme un ensemble de données dans lequel la soi-disant bonne réponse, c'est-à-dire l'étiquette ou le prix correct  $y$ , est donnée pour chaque maison de la parcelle. La tâche de l'algorithme d'apprentissage est de produire un plus grand nombre de ces bonnes réponses, en particulier en prédisant le prix probable d'autres maisons, comme la maison de votre ami. C'est pourquoi il s'agit d'un apprentissage supervisé. Pour définir un peu plus la terminologie, cette prévision du prix de l'immobilier correspond à un type particulier d'apprentissage supervisé appelé régression. Par régression, je veux dire que nous essayons de prédire un chiffre à partir d'une infinité de nombres possibles, tels que les prix de l'immobilier dans notre exemple, qui peuvent être de 150 000, 70 000 ou 183 000 ou tout autre chiffre intermédiaire. Il s'agit d'un apprentissage supervisé, d'entrées et de sorties d'apprentissage ou de mappages  $x$  à  $y$ .

## Apprentissage supervisé partie 2

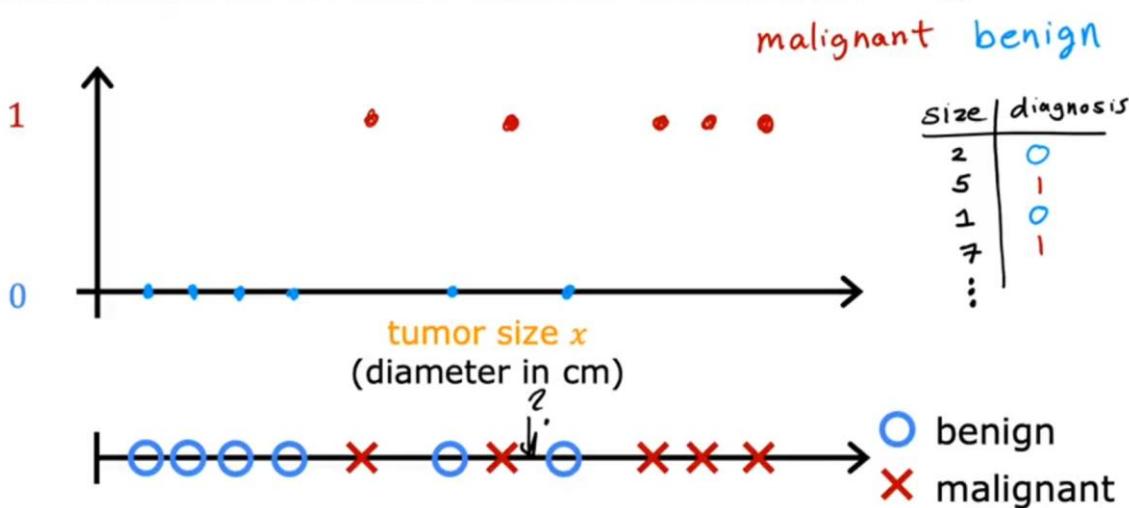
Prenons la détection du cancer du sein comme exemple d'un problème de classification. Supposons que vous développez un système d'apprentissage automatique afin que les médecins puissent disposer d'un outil de diagnostic pour détecter le cancer du sein. Ceci est important car une détection précoce peut potentiellement sauver la vie d'un patient. À l'aide du dossier médical d'un patient, votre système d'apprentissage automatique essaie de déterminer si une tumeur qui est une masse est maligne, c'est-à-dire cancéreuse ou dangereuse. Ou si cette tumeur, cette masse est bénigne, c'est-à-dire qu'il s'agit simplement d'une masse non cancéreuse et n'est-ce pas dangereux ?

## Classification: Breast cancer detection



Peut-être que votre ensemble de données contient des tumeurs de différentes tailles. Et ces tumeurs sont étiquetées soit comme bénignes, que je désignerai dans cet exemple par un 0, soit comme malignes, ce qui désignera dans cet exemple par un 1. Vous pouvez ensuite tracer vos données sur un graphique comme celui-ci où l'axe horizontal représente la taille de la tumeur et l'axe vertical ne prend que deux valeurs 0 ou 1 selon que la tumeur est bénigne, 0 ou maligne 1. L'une des raisons pour lesquelles cela est différent de la régression est que nous essayons de ne prévoir qu'un petit nombre de sorties ou de catégories possibles. Dans ce cas, deux sorties possibles 0 ou 1, bénignes ou malignes. Ceci est différent de la régression qui essaie de prédire n'importe quel nombre, l'ensemble du nombre infini de nombres possibles. C'est donc le fait qu'il n'y ait que deux sorties possibles qui fait cette classification.

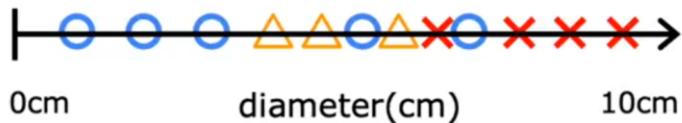
## Classification: Breast cancer detection



Comme il n'y a que deux sorties ou deux catégories possibles dans cet exemple, vous pouvez également tracer cet ensemble de données sur une ligne comme celle-ci. À l'heure actuelle, je vais utiliser deux symboles différents pour désigner la catégorie en utilisant un cercle, un O pour désigner les exemples bénins et une croix pour désigner les exemples malins. Et si de nouveaux patients se présentent pour un diagnostic et qu'ils présentent une masse de cette taille, alors la question est de savoir si votre système classera cette tumeur comme bénigne ou maligne ?

## Classification: Breast cancer detection

- benign
- ✗ malignant type 1
- △ malignant type 2

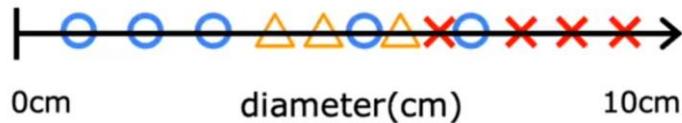


class category  
Classification

Il s'avère que dans les problèmes de classification, vous pouvez également avoir plus de deux catégories de sortie possibles. Peut-être que vous apprenez que l'algorithme peut générer plusieurs types de diagnostic de cancer s'il s'avère malin. Appelons donc deux types de cancer différents de type 1 et de type 2. Dans ce cas, la moyenne comporterait trois catégories de sortie possibles qu'elle pourrait prédire. D'ailleurs, dans la classification, les termes classes de sortie et catégories de sortie sont souvent utilisés de manière interchangeable. Donc, ce que je dis classe ou catégorie quand je fais référence à la sortie signifie la même chose.

## Classification: Breast cancer detection

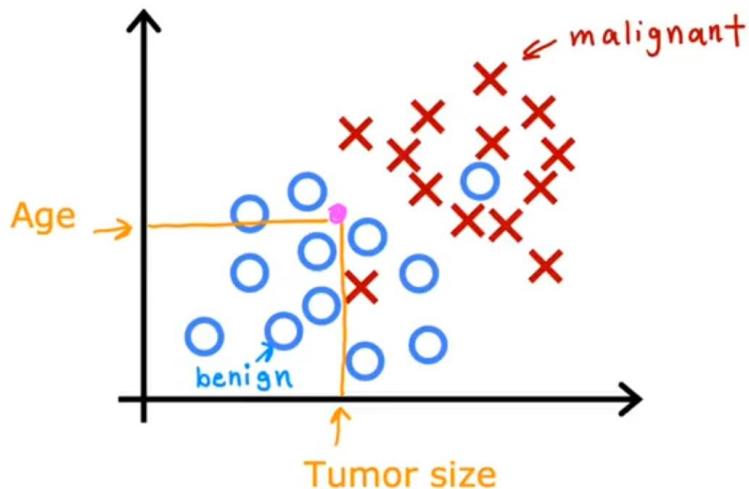
- benign
- ✗ malignant type 1
- △ malignant type 2



class category  
Classification  
predict categories cat dog benign malignant 0, 1, 2  
small number of possible outputs

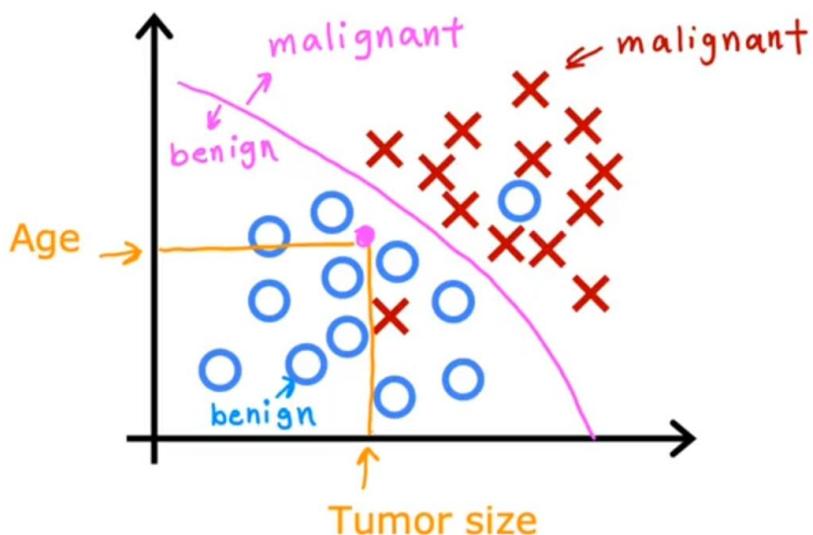
Donc, pour résumer, les algorithmes de classification prédisent les catégories. Les catégories ne doivent pas nécessairement être des chiffres. Il peut être non numérique, par exemple, il peut prédire si une image est celle d'un chat ou d'un chien. Et il peut prédire si une tumeur est bénigne ou maligne. Les catégories peuvent également être des nombres tels que 0, 1 ou 0, 1, 2. Mais ce qui différencie la classification de la régression lorsque vous interprétez les nombres, c'est que la classification prédit un petit ensemble limité de catégories de sortie possibles telles que 0, 1 et 2, mais pas tous les nombres possibles entre les deux, comme 0,5 ou 1,7.

## Two or more inputs



Dans l'exemple d'apprentissage supervisé que nous avons examiné, nous n'avions qu'une seule valeur d'entrée, la taille de la tumeur. Mais vous pouvez également utiliser plusieurs valeurs d'entrée pour prévoir une sortie. Voici un exemple, au lieu de simplement connaître la taille de la tumeur, supposons que vous ayez également l'âge de chaque patient en années. Votre nouvel ensemble de données comporte désormais deux entrées, l'âge et la taille de la tumeur. Dans ce nouvel ensemble de données, nous allons utiliser des cercles pour montrer les patients dont les tumeurs sont bénignes et des croix pour montrer les patients atteints d'une tumeur maligne. Ainsi, lorsqu'un nouveau patient arrive, le médecin peut mesurer la taille de la tumeur du patient et également enregistrer son âge. Dans ces conditions, comment pouvons-nous prédire si la tumeur de ce patient est bénigne ou maligne ? Eh bien, étant donné la journée ainsi décrite, l'algorithme d'apprentissage pourrait permettre de trouver une limite séparant les tumeurs malignes des tumeurs bénignes.

## Two or more inputs



L'algorithme d'apprentissage doit donc décider comment ajuster une ligne de démarcation à partir de ces données. La ligne de démarcation trouvée par l'algorithme d'apprentissage aiderait le médecin à établir le diagnostic. Dans ce cas, la tumeur est plus susceptible d'être bénigne. À partir de cet exemple, nous avons vu comment saisir l'âge du patient et la taille de la tumeur peuvent être utilisés.

Dans d'autres problèmes d'apprentissage automatique, beaucoup plus de valeurs d'entrée sont souvent requises.

## Supervised learning

Learns from being given “right answers”

### Regression

Predict a **number**

**infinitely** many possible outputs

### Classification

**predict categories**

**small number** of possible outputs

## Résumé

Ainsi, pour récapituler, l'apprentissage supervisé fait correspondre l'entrée  $x$  à la sortie  $y$ , où l'algorithme d'apprentissage apprend à partir des bonnes réponses entre guillemets. Les deux principaux types d'apprentissage supervisé sont la régression et la classification. Dans une application de régression telle que la prédiction des prix des maisons, l'algorithme d'apprentissage doit prédire les nombres à partir d'une infinité de nombres de sortie possibles. Alors qu'en classification, l'algorithme d'apprentissage doit faire une prédiction d'une catégorie, le tout à partir d'un petit ensemble de résultats possibles. Vous savez donc maintenant ce qu'est l'apprentissage supervisé, y compris la régression et la classification.