Artificial Machines Learning Intelligence

Exploring data and identifying patterns, and involves minimal human intervention.

Scenario: Data for a bicycle sharing scheme was collected and used to predict the number of rentals based on seasonality and weather conditions.

Artificial Intelligence is the ability of computers/machines to perform tasks that usually people perform. For example, the ability to learn, make decisions, and solve problems.

AI works via special software and includes algorithms and technologies (such as supervised and unsupervised machine learning) that allow machines to learn from experience and adjust their actions without human intervention.

The purpose of artificial intelligence is to take actions that have the best chance of achieving a specific goal.

Symbolic AI, neural networks, expert systems, machine perception, conversational natural language processing, adaptive motor control and cognitive architecture.

Parameter values are not selected by the human designer. Instead, parameter values are set to an initial guess, and then adjusted during an automated learning process called training.

Models don't train themselves; they're trained using data plus two pieces of code: the objective function, and the optimizer.

Training only changes the parameter values inside of a model—it doesn't change what kind of model is used.

- **The Goal** of training the model is to find a function that performs some kind of calculation to the x values that produces the result y.
- **How**: By applying a machine learning algorithm that tries to fit the x values to a calculation that produces y reasonably accurately for all of the cases in the training dataset.

Regression works by establishing a relationship between variables in the data that represent characteristics—known as the features—of the thing being observed, and the variable we're trying to predict—known as the label.

- Consider the company that rents bicycles and wants to **predict** the expected number of rentals in a given day.

In this case, features include things like the day of the week, month, and so on, while the label is the number of bicycle rentals.

- **Features** (dimensions) include *things* like the day of the week, month, and so on
- **Labels** (factTables) are the *number* of bicycle rentals.

The use of historic data with known label values to train a model makes regression an example of supervised machine learning.

- A *training dataset* to which we'll apply an algorithm that determines a function encapsulating the relationship between the feature values and the known label values.
- A *validation or test dataset* that we can use to evaluate the model by using it to generate predictions for the label and comparing them to the actual known label values.

Understand the relationships between its attributes; in particular, any apparent correlation between the features (*things*) and the label (*numbers*) your model will try to predict.

Determine x and y: Rentals represents the label (the y value) our model must be trained to predict. The other columns are potential features (x values)

1. To detect and fix issues in the data:
    a. Missing values
    b. Errors
    c. Outlier values
    d. Deriving new feature columns by transforming or combining existing features (a process known as feature engineering)
    e. Normalizing numeric features (values you can measure or count) so they're on a similar scale
    f. Encoding categorical features (values that represent discrete categories) as numeric indicators.
2. Space
3. Clearing the confusion: fig, ax = plt.subplots() | Towards Data Science
4. fig, ax = plt.subplots()
5. Created two variables, fig and ax
6. These two variables now hold the two core objects used for all types of plotting operations.
    a. First object fig, short for figure, imagine it as the frame of your plot. You can resize, reshape the frame but you cannot draw on it. On a single notebook or a script, you can have multiple figures. Each figure can have multiple subplots.
    b. The second object, ax, short for axes, is the canvas you draw on. Or rephrasing, it is the blank sheet you can plot and hold your data. An axes object can only belong to one figure.
    c. Clearing the confusion: fig, ax = plt.subplots() | Towards Data Science
7. space

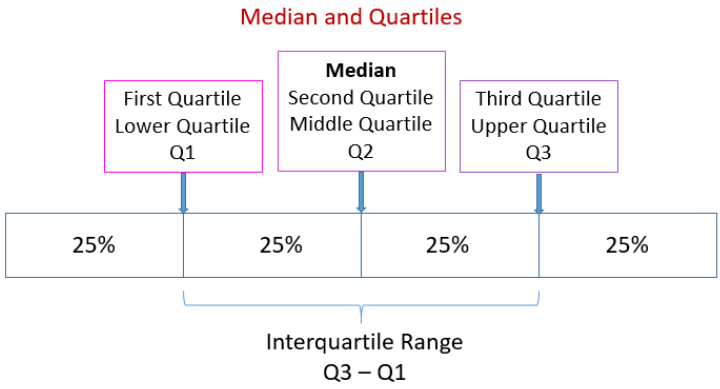
Train a Regression Model

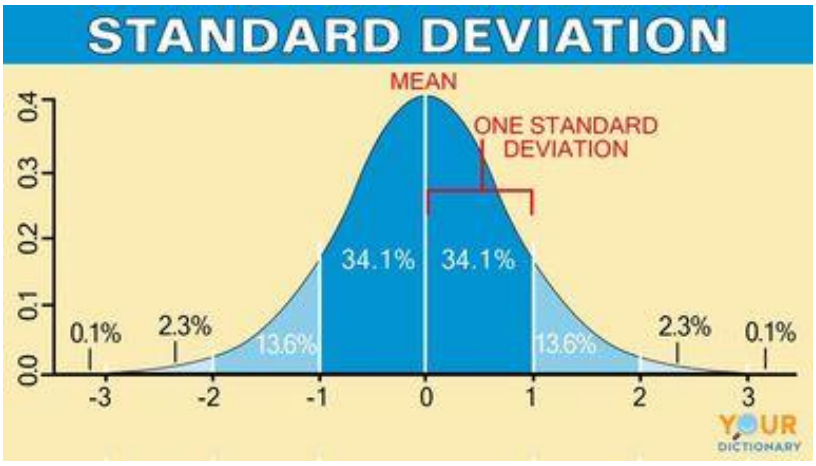
1. Separate the features we want to use to train the model from the label we want it to predict.
2. Now have NumPy arrays named X containing the features, and y containing the labels.
3. Randomly split the data into two subsets; train_test_split function;
    a. A (typically larger) set with which to train the model
    b. A smaller "hold-back" set with which to validate the trained model.
4. Train a model by fitting a suitable regression algorithm to the training data.
5. Simple models with small datasets can often be fit in a single step, while larger datasets and more complex models must be fit by repeatedly using the model with training data and

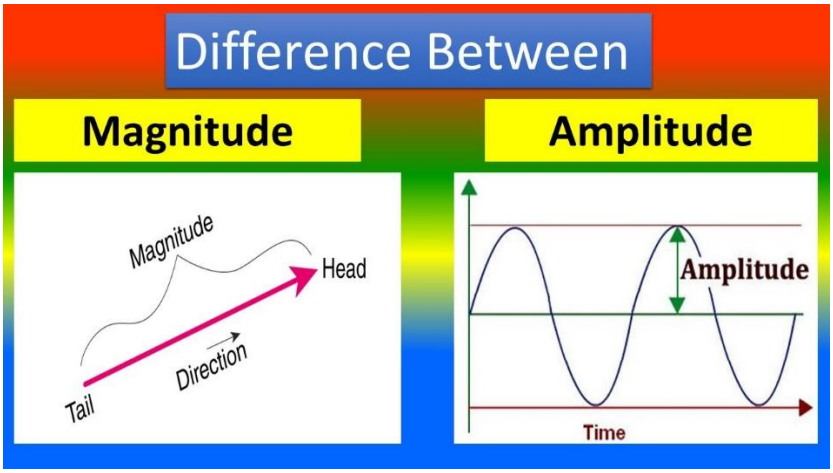
comparing the output with the expected label. If the prediction is accurate enough, we consider the model trained. If not, we adjust the model slightly and loop again.

6. Hyperparameters are values that change the way that the model is fit during these loops.
7. Preprocessing refers to changes you make to your data before it is passed to the model.
8. Scale features so they fall between zero and one.
9. Use categorical features such as 'bicycle', 'skateboard' or 'car'. (*Make everything a number*)
    a. These features are represented by 0 or 1 values in one-hot vectors
    b. Vectors that have a 0 or 1 for each possible value.
    c. For example, bicycle, skateboard, and car might respectively be (1,0,0), (0,1,0), and (0,0,1).
    d. Apply ordinal encoding to substitute a unique integer value for each category?
    e. Appy one hot encoding to create individual binary (0 or 1) features for each possible category
    f. Pipelines: Enable defining a set of preprocessing steps that end with an algorithm.
10. space

[Median, Quartiles, Percentiles (video lessons, examples, solutions) (onlinemathlearning.com)]







space

Solving Word Questions (mathsisfun.com)

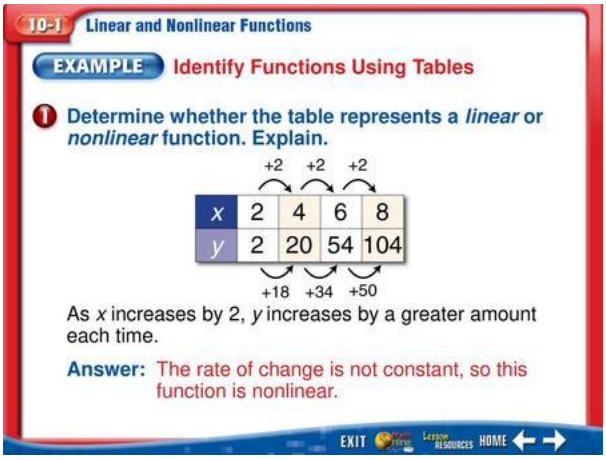
How to use a model - Learn | Microsoft Docs

System Data | Capital Bikeshare – shared data for labs

1. Load data from file, filter it, and graph it.
2. printf is short for print formatter, it gives you the ability to mark where in the String variables will go and pass in those variables with it. This saves from having to do a long String concatenation.
3. A dataset is a collection of information about objects or things.
4. The goal of machine learning is to find patterns in data and use this to make estimates
5. Machine learning differs from normal software development in that we use special code, rather than our own intuition, to improve how well the software works.
6. The learning process conceptually uses four components:
    a. Data about the topic we are interested in.
    b. A model, which makes estimates.
    c. An objective the model is trying to achieve.
    d. An optimizer, which is the additional code that changes the model depending on its performance.
7. Data can be thought of as features, and labels. Features correspond to potential model inputs, while labels correspond to model outputs, or desired model outputs
8. Pandas and Plotly are powerful tools to explore datasets in Python
9. Once we have a trained model, we can save to disk for later use.

Machine Learning: *"Artificial Machines Learning Intelligence"*

1. Azure Machine Learning designer | Microsoft Azure
2. What is the Azure Machine Learning designer? - Azure Machine Learning | Microsoft Docs
3. model selection - How to choose the number of hidden layers and nodes in a feedforward neural network? - Cross Validated (stackexchange.com)
4. Create a Training Job with the job creation UI - Azure Machine Learning | Microsoft Docs
5. Tutorial: Designer - train a no-code regression model - Azure Machine Learning | Microsoft Docs
6. Tutorial data: Don't Get Kicked! | Kaggle
    a. Inside Kaggle you'll find all the code & data you need to do your data science work.
7. Tutorial: Create Training and Inferencing Pipelines with Azure ML Designer – The New Stack
8. Tutorial: Designer - deploy no-code models - Azure Machine Learning | Microsoft Docs
9. Run Real-time Predictions with an Inference Pipeline - Amazon SageMaker
10. Azure Machine Learning REST APIs | Microsoft Docs
11.
12. https://docs.microsoft.com/en-us/office/troubleshoot/access/database-normalization-description
13.
14. https://azure.microsoft.com/en-us/services/machine-learning/designer/
15.
16. https://docs.microsoft.com/en-us/azure/machine-learning/tutorial-designer-automobile-price-train-score
17.
18. https://thenewstack.io/tutorial-create-training-and-inferencing-pipelines-with-azure-ml-designer/
19. Space



Terminology

1. *MLOps* or ML Ops is a set of practices that aims to deploy and maintain machine learning models in production reliably and efficiently. The word is a compound of "machine learning" and the continuous development practice of DevOps in the software field.
2. *Featurization* is the process to convert varied forms of data to numerical data which can be used for basic ML algorithms. Data can be text data, images, videos, graphs, various database tables, time-series, categorical features, etc.
3. *Training data* is the data you use to train an algorithm or machine learning model to predict the outcome you design your model to predict. Test data is used to measure the performance, such as accuracy or efficiency, of the algorithm you are using to train the machine.
4. Why we use OLS model?
5. In data analysis, we use OLS for estimating the unknown parameters in a linear regression model. The goal is minimizing the differences between the collected observations in some arbitrary dataset and the responses predicted by the linear approximation of the data.
6. space

Each of these stages maps to a set of modules in Azure ML Studio.

1. Step 1: preprocess your data. For this step, you'll be using the modules under Data Transformation . ...
2. Step 2: split into train/test. ...
3. Step 3: select and/or create your features. ...
4. Step 4–5–6: train, score and evaluate your model. ...
5. Step 7: deploy selected model.

References

1. Introduction to machine learning - Learn | Microsoft Docs
2. Build classical machine learning models with supervised learning - Learn | Microsoft Docs
3. Foundations of data science for machine learning - Learn | Microsoft Docs
4. Understand data science for machine learning - Learn | Microsoft Docs
5. Space