1. Introduction to Machine Learning

Shabana K M
PhD Research Scholar
Computer Science and Engineering

IIT Palakkad

31 July 2021

INDIAN INSTITUTE OF TECHNOLOGY PALAKKAD

IEEE STUDENT BRANCH GEC PALAKKAD IEEE LINK INTERNZONE IEEE Kerala Section IEEE

Artificial Intelligence (AI)

Artificial Intelligence (AI)

- intelligence demonstrated by machines

Artificial Intelligence (AI)

- intelligence demonstrated by machines
- AI systems develop the ability to rationalize and perform actions that have the best chance of achieving a specific goal

Artificial Intelligence (AI)

- intelligence demonstrated by machines
- AI systems develop the ability to rationalize and perform actions that have the best chance of achieving a specific goal
- traditional AI problems include - reasoning, planning, learning, natural language processing

Artificial Intelligence (AI)

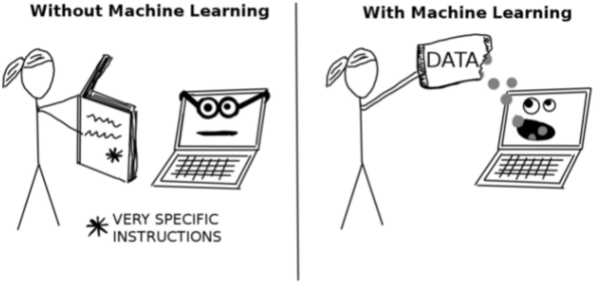
- intelligence demonstrated by machines
- AI systems develop the ability to rationalize and perform actions that have the best chance of achieving a specific goal
- traditional AI problems include - reasoning, planning, learning, natural language processing
- autonomous vehicles, playing games, search engines, online assistants, image recognition in photographs, spam filtering

Machine Learning

study of computer algorithms that improve automatically through experience

Machine Learning

study of computer algorithms that improve automatically through experience



Machine Learning

- "A computer program is said to **learn** from experience `E` with respect to some class of tasks `T` and performance measure `P`, if its performance at tasks in `T`, as measured by `P`, improves with experience `E` " - `Tom Mitchell`

Machine Learning

- "A computer program is said to **learn** from experience `E` with respect to some class of tasks `T` and performance measure `P`, if its performance at tasks in `T`, as measured by `P`, improves with experience `E` " - `Tom Mitchell`
- **Example:** Classifying e-mails as spam or not

Machine Learning

- "A computer program is said to **learn** from experience `E` with respect to some class of tasks `T` and performance measure `P`, if its performance at tasks in `T`, as measured by `P`, improves with experience `E` " - `Tom Mitchell`
- **Example:** Classifying e-mails as spam or not
  Here,
  - `T` - predict whether an email is spam or not

Machine Learning

- "A computer program is said to **learn** from experience `E` with respect to some class of tasks `T` and performance measure `P`, if its performance at tasks in `T`, as measured by `P`, improves with experience `E` " - `Tom Mitchell`
- **Example:** Classifying e-mails as spam or not
  Here,
  - `T` - predict whether an email is spam or not
  - `E` - collection of spam and non-spam mails

Machine Learning

- "A computer program is said to **learn** from experience `E` with respect to some class of tasks `T` and performance measure `P`, if its performance at tasks in `T`, as measured by `P`, improves with experience `E` " - `Tom Mitchell`
- **Example:** Classifying e-mails as spam or not
  Here,
  - `T` - predict whether an email is spam or not
  - `E` - collection of spam and non-spam mails
  - `P` - (i): *accuracy of prediction* - percentage of correct predictions

Machine Learning

- "A computer program is said to **learn** from experience `E` with respect to some class of tasks `T` and performance measure `P`, if its performance at tasks in `T`, as measured by `P`, improves with experience `E` " - `Tom Mitchell`
- **Example:** Classifying e-mails as spam or not
  Here,
  - `T` - predict whether an email is spam or not
  - `E` - collection of spam and non-spam mails
  - `P` - (i): *accuracy of prediction* - percentage of correct predictions
    (ii): *precision* - proportion of true spam mails among those classified as spam

Datasets for machine learning

Attributes

sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data point /example

Numerical value

Categorical value

Dataset

Datasets for machine learning

Attributes

sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data point /example

Numerical value

Categorical value

Dataset

- table with the data from which a machine learns

Datasets for machine learning

Attributes

Data point /example

Numerical value

Categorical value

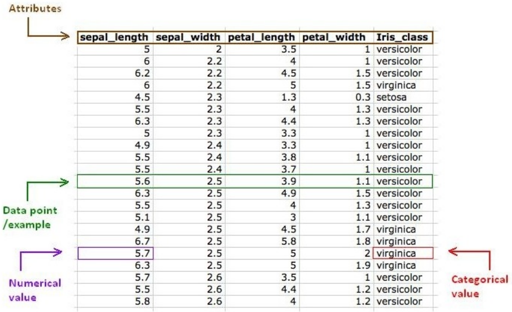
sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Dataset

- table with the data from which a machine learns
- contains features (columns) and observations (rows)

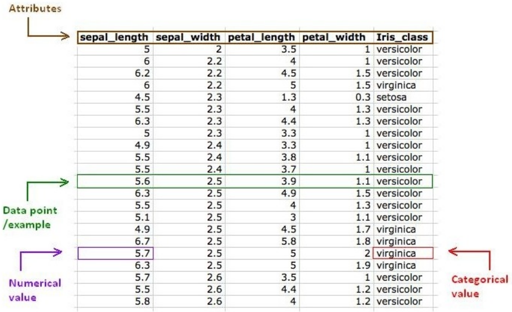
Datasets for machine learning

Features/Attributes



sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Datasets for machine learning

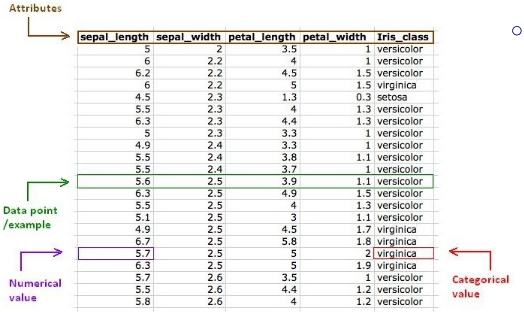


sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Features/Attributes

- columns in the dataset

Datasets for machine learning



sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Features/Attributes

- columns in the dataset
- describes data of a single type

Datasets for machine learning

Attributes

Data point /example

Numerical value

Categorical value

sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Features/Attributes

- columns in the dataset
- describes data of a single type
- can be Qualitative (categorical) or Quantitative (numerical)

Datasets for machine learning

Attributes

sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data point /example

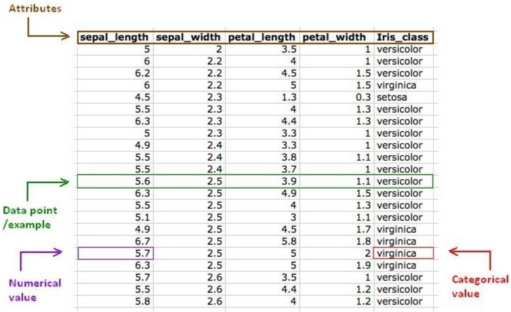
Numerical value

Categorical value

Features/Attributes

- columns in the dataset
- describes data of a single type
- can be Qualitative (categorical) or Quantitative (numerical)
- gender of a person, grades obtained in a test, etc. are qualitative

Datasets for machine learning



sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Features/Attributes

- columns in the dataset
- describes data of a single type
- can be Qualitative (categorical) or Quantitative (numerical)
- gender of a person, grades obtained in a test, etc. are qualitative
- height, weight, temperature, etc. are quantitative

Datasets for machine learning

Data points/Observations

Attributes

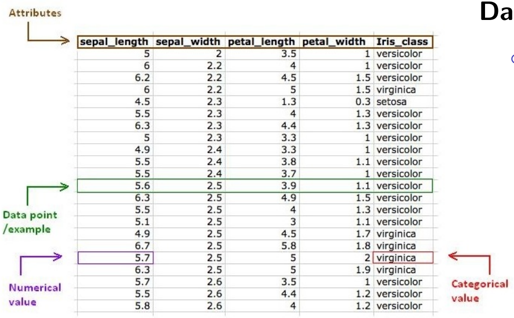
sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data point /example

Numerical value

Categorical value

Datasets for machine learning

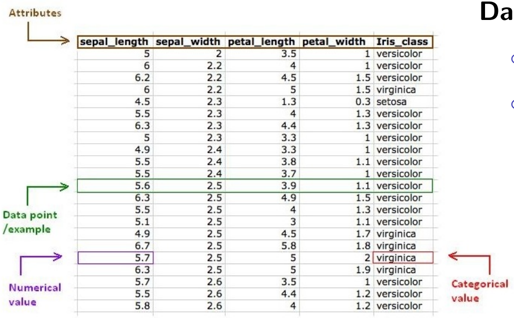


sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data points/Observations

- rows in the dataset

Datasets for machine learning



sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data points/Observations

- rows in the dataset
- also called instance or example

Datasets for machine learning

Attributes

sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data point /example

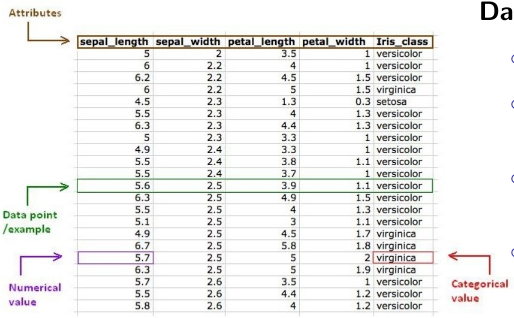
Numerical value

Categorical value

Data points/Observations

- rows in the dataset
- also called instance or example
- describes a single entity or observation

Datasets for machine learning



sepal_length	sepal_width	petal_length	petal_width	Iris_class
5	2	3.5	1	versicolor
6	2.2	4	1	versicolor
6.2	2.2	4.5	1.5	versicolor
6	2.2	5	1.5	virginica
4.5	2.3	1.3	0.3	setosa
5.5	2.3	4	1.3	versicolor
6.3	2.3	4.4	1.3	versicolor
5	2.3	3.3	1	versicolor
4.9	2.4	3.3	1	versicolor
5.5	2.4	3.8	1.1	versicolor
5.5	2.4	3.7	1	versicolor
5.6	2.5	3.9	1.1	versicolor
6.3	2.5	4.9	1.5	versicolor
5.5	2.5	4	1.3	versicolor
5.1	2.5	3	1.1	versicolor
4.9	2.5	4.5	1.7	virginica
6.7	2.5	5.8	1.8	virginica
5.7	2.5	5	2	virginica
6.3	2.5	5	1.9	virginica
5.7	2.6	3.5	1	versicolor
5.5	2.6	4.4	1.2	versicolor
5.8	2.6	4	1.2	versicolor

Data points/Observations

- rows in the dataset
- also called instance or example
- describes a single entity or observation
- features describe properties about that observation

Some basic terminology

Algorithm

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner
- procedure that is run on data to create a machine learning model

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner
- procedure that is run on data to create a machine learning model

Machine learning model

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner
- procedure that is run on data to create a machine learning model

Machine learning model

- output from a machine learning algorithm

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner
- procedure that is run on data to create a machine learning model

Machine learning model

- output from a machine learning algorithm
- represents what was learned by a machine learning algorithm

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner
- procedure that is run on data to create a machine learning model

Machine learning model

- output from a machine learning algorithm
- represents what was learned by a machine learning algorithm
- learned program that maps inputs to predictions

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner
- procedure that is run on data to create a machine learning model

Machine learning model

- output from a machine learning algorithm
- represents what was learned by a machine learning algorithm
- learned program that maps inputs to predictions

Target or Label

Some basic terminology

Algorithm

- set of rules that a machine follows to achieve a particular goal
- can be considered as a recipe that defines the inputs, the output and all the steps needed to get from the inputs to the output

Machine learning algorithm

- also called learner
- procedure that is run on data to create a machine learning model

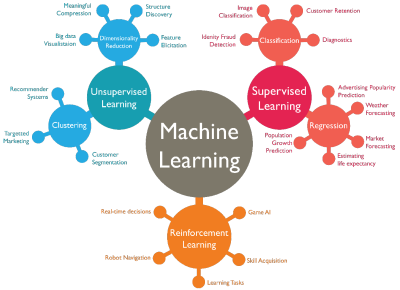
Machine learning model

- output from a machine learning algorithm
- represents what was learned by a machine learning algorithm
- learned program that maps inputs to predictions

Target or Label

- the value predicted by a machine learning model

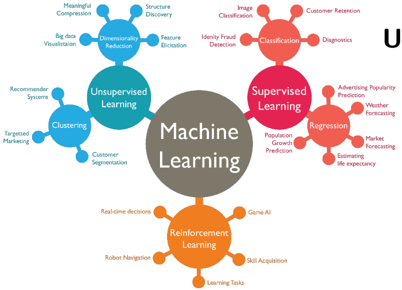
Types of machine learning algorithms



Supervised learning

- work on data sets that include its desired outputs (or labels)

Types of machine learning algorithms



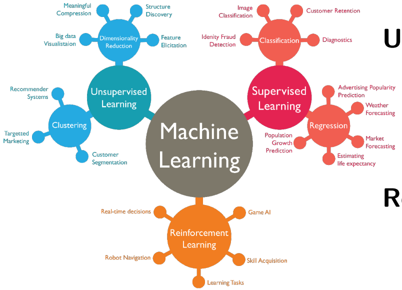
Supervised learning

- work on data sets that include its desired outputs (or labels)

Unsupervised learning

- these algorithms detect patterns, mine rules and group data points without labels

Types of machine learning algorithms



Supervised learning

- work on data sets that include its desired outputs (or labels)

Unsupervised learning

- these algorithms detect patterns, mine rules and group data points without labels

Reinforcement learning

- algorithms attempt to learn actions that would maximize the reward by continuously interacting with the environment

Supervised learning

Supervised learning problems are categorized into:

Supervised learning

Supervised learning problems are categorized into:

- regression
- classification

Supervised learning

Supervised learning problems are categorized into:

- regression
- classification

Regression

Target value is continuous
`Examples:` forecasting stock prices, predicting house prices, etc.

Supervised learning

Supervised learning problems are categorized into:

- regression
- classification

Regression

Target value is continuous
`Examples:` forecasting stock prices, predicting house prices, etc.

Classification

Output is discrete

Supervised learning

Supervised learning problems are categorized into:

- regression
- classification

Regression

Target value is continuous
`Examples:` forecasting stock prices, predicting house prices, etc.

Classification

Output is discrete

- *Binary classification* - only two classes
  eg:- spam filtering, fraud detection, etc.

Supervised learning

Supervised learning problems are categorized into:

- regression
- classification

Regression

Target value is continuous
`Examples:` forecasting stock prices, predicting house prices, etc.

Classification

Output is discrete

- *Binary classification* - only two classes
  eg:- spam filtering, fraud detection, etc.
- *Multi-class classification* - three or more classes
  eg:- recognition of handwritten characters, object detection, etc.

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture   Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles Unsupervised

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles Unsupervised
4. Identify the numbers in a handwritten ZIP code, from a digitized image

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles Unsupervised
4. Identify the numbers in a handwritten ZIP code, from a digitized image Classification

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles Unsupervised
4. Identify the numbers in a handwritten ZIP code, from a digitized image Classification
5. Recommend movies based on user ratings

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles Unsupervised
4. Identify the numbers in a handwritten ZIP code, from a digitized image Classification
5. Recommend movies based on user ratings Unsupervised

Categorize these learning problems!

1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles Unsupervised
4. Identify the numbers in a handwritten ZIP code, from a digitized image Classification
5. Recommend movies based on user ratings Unsupervised
6. Personalized recommendation of videos/music

Categorize these learning problems!

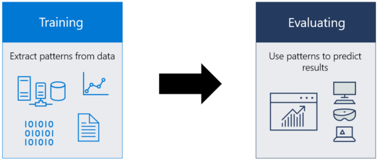
1. Given a picture of a male/female, predict his/her age based on given picture Regression
2. For a bank, decide whether or not to give a loan to someone on the basis of his credit history Classification
3. Discover trends in news articles Unsupervised
4. Identify the numbers in a handwritten ZIP code, from a digitized image Classification
5. Recommend movies based on user ratings Unsupervised
6. Personalized recommendation of videos/music Reinforcement learning

Demystifying “a Machine Learning Model”

In supervised learning, a dataset is comprised of input features and a target variable

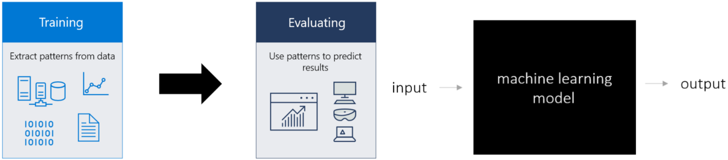
Demystifying "a Machine Learning Model"

In supervised learning, a dataset is comprised of input features and a target variable



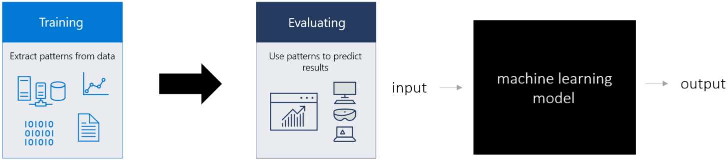
Demystifying “a Machine Learning Model”

In supervised learning, a dataset is comprised of input features and a target variable



Demystifying "a Machine Learning Model"

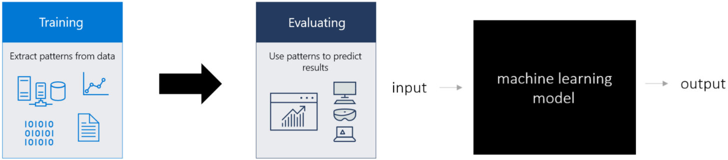
In supervised learning, a dataset is comprised of input features and a target variable



We assume that there is an unknown underlying function that maps the input features to the target variable

Demystifying "a Machine Learning Model"

In supervised learning, a dataset is comprised of input features and a target variable

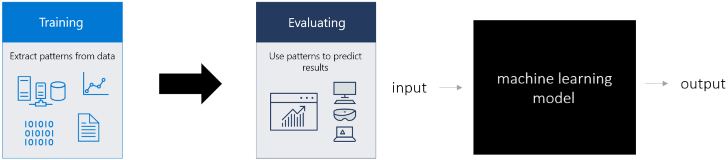


We assume that there is an unknown underlying function that maps the input features to the target variable

- This true function is referred to as the target function

Demystifying “a Machine Learning Model”

In supervised learning, a dataset is comprised of input features and a target variable

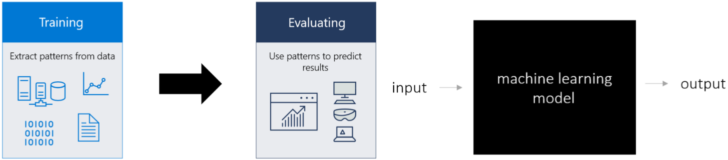


We assume that there is an unknown underlying function that maps the input features to the target variable

- This true function is referred to as the target function
- A supervised learning algorithm tries to estimate the target function using historical or available observations

Demystifying "a Machine Learning Model"

In supervised learning, a dataset is comprised of input features and a target variable



We assume that there is an unknown underlying function that maps the input features to the target variable

- This true function is referred to as the target function
- A supervised learning algorithm tries to estimate the target function using historical or available observations – Function approximation!

Demystifying “a Machine Learning Model”

Machine learning model

Given a set of input features $x = x_1, x_2, ..., x_n$, machine learning algorithms learn a target function f that best maps the input features (x) to an output variable (y)

$$y = f(x)$$

Demystifying “a Machine Learning Model”

Machine learning model

Given a set of input features $x = x_1, x_2, ..., x_n$, machine learning algorithms learn a target function f that best maps the input features (x) to an output variable (y)

$$y = f(x)$$

`The learned function` f `along with its parameters form the ML model`

Demystifying "a Machine Learning Model"

Machine learning model

Given a set of input features $x = x_1, x_2, ..., x_n$, machine learning algorithms learn a target function f that best maps the input features (x) to an output variable (y)

$$y = f(x)$$

`The learned function` f `along with its parameters form the ML model`

Function parameters

- quantities with a constant value that influence the output of a function

Demystifying "a Machine Learning Model"

Machine learning model

Given a set of input features $x = x_1, x_2, ..., x_n$, machine learning algorithms learn a target function f that best maps the input features (x) to an output variable (y)

$$y = f(x)$$

`The learned function` f `along with its parameters form the ML model`

Function parameters

- quantities with a constant value that influence the output of a function
- for instance, consider the function $f(x) = ax^2 + bx + c$. Here x is considered as the input and $a, b\ and\ c$ are the parameters

Demystifying "a Machine Learning Model"

Machine learning model

Given a set of input features $x = x_1, x_2, ..., x_n$, machine learning algorithms learn a target function f that best maps the input features (x) to an output variable (y)

$$y = f(x)$$

`The learned function` f `along with its parameters form the ML model`

Function parameters

- quantities with a constant value that influence the output of a function
- for instance, consider the function $f(x) = ax^2 + bx + c$. Here x is considered as the input and $a, b \; and \; c$ are the parameters
- for each set of parameter values, we get a different function

Demystifying “a Machine Learning Model”

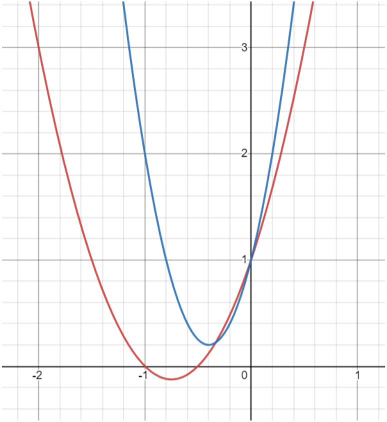


Figure: Plotting the functions $2x^2 + 3x + 1$(red) and $5x^2 + 4x + 1$(blue)

Demystifying “a Machine Learning Model”

- Different machine learning algorithms make different assumptions about the form of the function being learned (eg;- linear, nonlinear, etc.) and how best to approximate it

Demystifying “a Machine Learning Model”

- Different machine learning algorithms make different assumptions about the form of the function being learned (eg;- linear, nonlinear, etc.) and how best to approximate it
- The learning algorithm estimates or learns the function parameters (model parameters) from the data

Demystifying “a Machine Learning Model”

- Different machine learning algorithms make different assumptions about the form of the function being learned (eg;- linear, nonlinear, etc.) and how best to approximate it
- The learning algorithm estimates or learns the function parameters (model parameters) from the data
- This is achieved by calculating the error between the `predicted outputs` and the `expected outputs` and adjusting the parameter values such that this error is minimized

Demystifying “a Machine Learning Model”

- Different machine learning algorithms make different assumptions about the form of the function being learned (eg;- linear, nonlinear, etc.) and how best to approximate it
- The learning algorithm estimates or learns the function parameters (model parameters) from the data
- This is achieved by calculating the error between the `predicted outputs` and the `expected outputs` and adjusting the parameter values such that this error is minimized
- The learned model can then be used for making predictions for unseen observations

Demystifying “a Machine Learning Model”

- Different machine learning algorithms make different assumptions about the form of the function being learned (eg;- linear, nonlinear, etc.) and how best to approximate it
- The learning algorithm estimates or learns the function parameters (model parameters) from the data
- This is achieved by calculating the error between the `predicted outputs` and the `expected outputs` and adjusting the parameter values such that this error is minimized
- The learned model can then be used for making predictions for unseen observations
- The more the data we have, the better the approximation of the target function

Demystifying "a Machine Learning Model"

- Approximating the target function is a hard task!!

Demystifying "a Machine Learning Model"

- Approximating the target function is a hard task!!
- In reality, there is an error (e) that is independent of the input data x such that $y = f(x) + e$

Demystifying “a Machine Learning Model”

- Approximating the target function is a hard task!!
- In reality, there is an error (e) that is independent of the input data x such that $y = f(x) + e$
- This error might be due to not having enough features to sufficiently characterize the best mapping from the given input x to the target variable y

Demystifying “a Machine Learning Model”

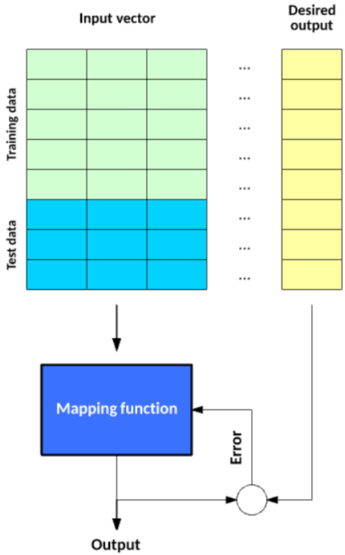
- Approximating the target function is a hard task!!
- In reality, there is an error (e) that is independent of the input data x such that $y = f(x) + e$
- This error might be due to not having enough features to sufficiently characterize the best mapping from the given input x to the target variable y
- e is is called irreducible error because no matter how good we get at estimating the target function (f), we cannot reduce this error

Demystifying "a Machine Learning Model"

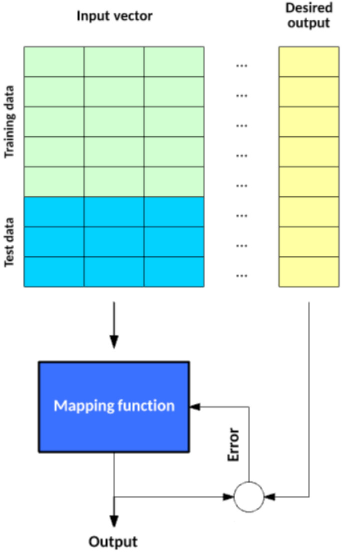
- Approximating the target function is a hard task!!
- In reality, there is an error (e) that is independent of the input data x such that $y = f(x) + e$
- This error might be due to not having enough features to sufficiently characterize the best mapping from the given input x to the target variable y
- e is is called irreducible error because no matter how good we get at estimating the target function (f), we cannot reduce this error
- Approximating the target function could also become difficult due to the presence of noisy features

Training and Test datasets

Training and Test datasets

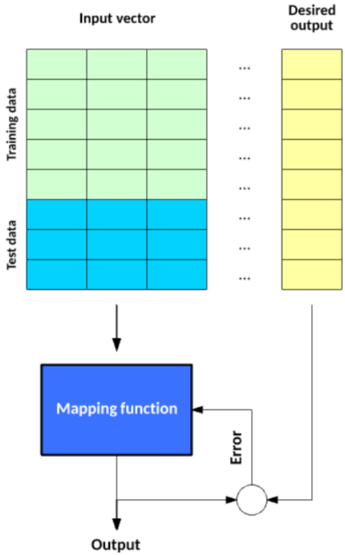


Training and Test datasets



Dataset is segmented into two types of samples:

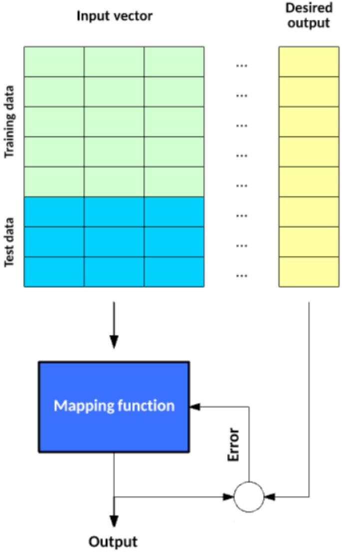
Training and Test datasets



Dataset is segmented into two types of samples:

- training data
- test data

Training and Test datasets

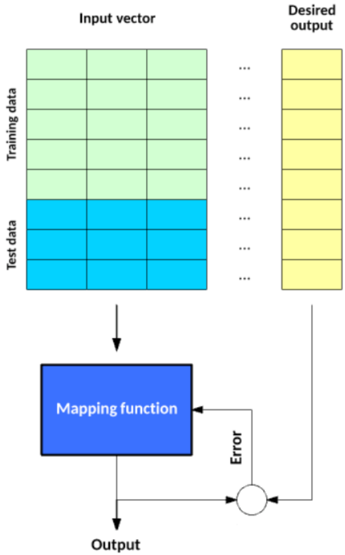


Dataset is segmented into two types of samples:

- training data
- test data

A model is learned using the examples in the training data

Training and Test datasets



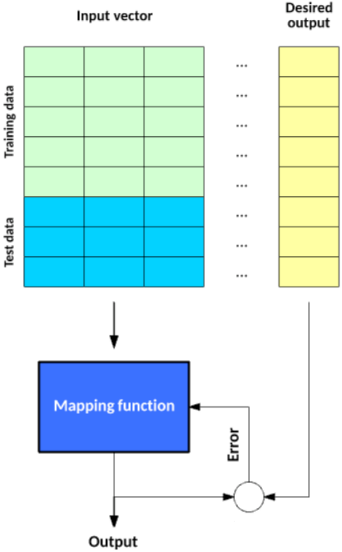
Dataset is segmented into two types of samples:

- training data
- test data

A model is learned using the examples in the training data

The trained model is then applied to test data

Training and Test datasets



Dataset is segmented into two types of samples:

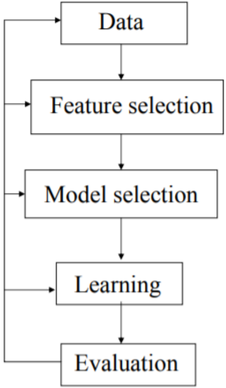
- training data
- test data

A model is learned using the examples in the training data

The trained model is then applied to test data

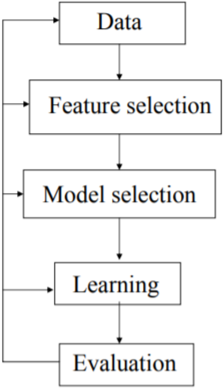
Performance on test data provides a good measure of how well the learned model generalizes to unseen data

Designing a machine learning system



Collecting Data

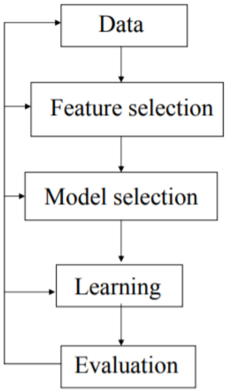
Designing a machine learning system



Collecting Data

- data can be
  - `structured` - data stored in rows and columns, such as spreadsheets

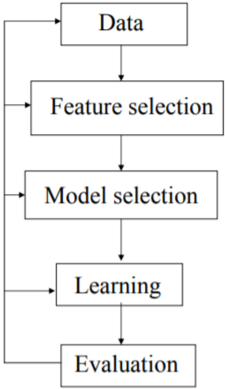
Designing a machine learning system



Collecting Data

- data can be
  - `structured` - data stored in rows and columns, such as spreadsheets
  - `unstructured` - text, audio, images, etc.

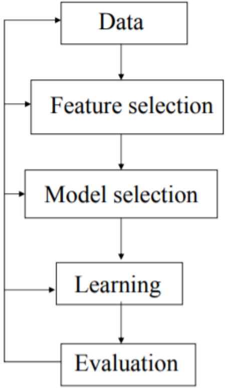
Designing a machine learning system



Collecting Data

- data can be
  - `structured` - data stored in rows and columns, such as spreadsheets
  - `unstructured` - text, audio, images, etc.
  - `semi-structured` - emails

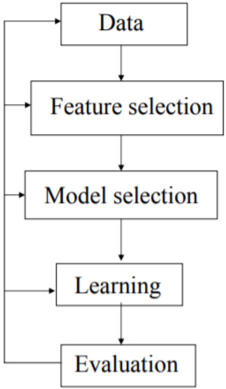
Designing a machine learning system



Collecting Data

- data can be
  - `structured` - data stored in rows and columns, such as spreadsheets
  - `unstructured` - text, audio, images, etc.
  - `semi-structured` - emails
- *the more the better!*

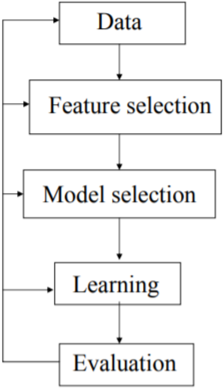
Designing a machine learning system



Collecting Data

- data can be
  - `structured` - data stored in rows and columns, such as spreadsheets
  - `unstructured` - text, audio, images, etc.
  - `semi-structured` - emails
- *the more the better!*
- data usually requires
  - `cleaning` - fixing incorrect data, dealing with missing values, etc.

Designing a machine learning system

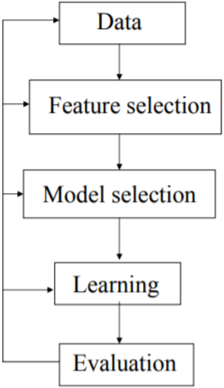


Collecting Data

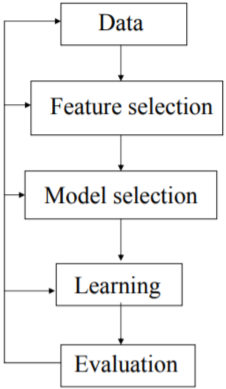
- data can be
  - `structured` - data stored in rows and columns, such as spreadsheets
  - `unstructured` - text, audio, images, etc.
  - `semi-structured` - emails
- *the more the better!*
- data usually requires
  - `cleaning` - fixing incorrect data, dealing with missing values, etc.
  - `preprocessing` - rescaling, discretization, etc.

Designing a machine learning system

Features



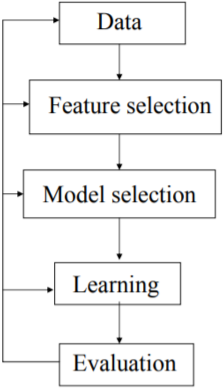
Designing a machine learning system



Features

- also called attributes or variables
  - measurable pieces of data used for analysis
  - appear as columns in structured data sets

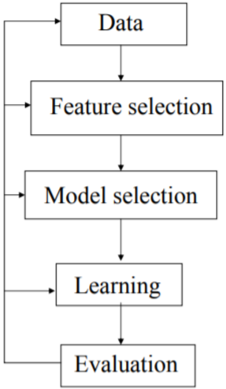
Designing a machine learning system



Features

- also called attributes or variables
  - measurable pieces of data used for analysis
  - appear as columns in structured data sets
- quality of features have a major impact on model performance

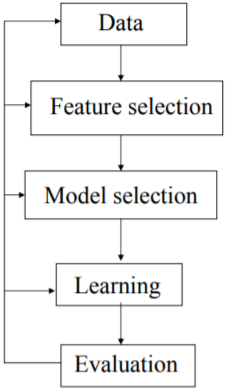
Designing a machine learning system



Features

- also called attributes or variables
  - measurable pieces of data used for analysis
  - appear as columns in structured data sets
- quality of features have a major impact on model performance
- `feature engineering` - generate new features from existing ones (eg:- BMI from height and weight)

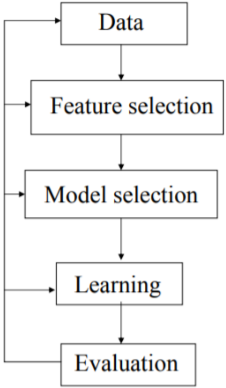
Designing a machine learning system



Features

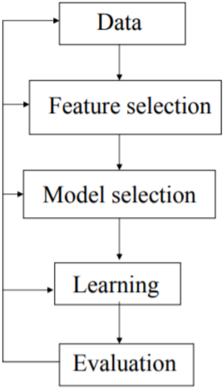
- also called attributes or variables
  - measurable pieces of data used for analysis
  - appear as columns in structured data sets
- quality of features have a major impact on model performance
- `feature engineering` - generate new features from existing ones (eg:- BMI from height and weight)
- `feature selection` - choose features contributing the most to model performance

Designing a machine learning system



Model Selection

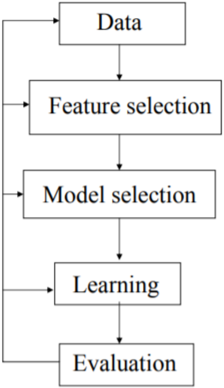
Designing a machine learning system



Model Selection

- selecting the best model from among a collection of candidate machine learning models for the task at hand

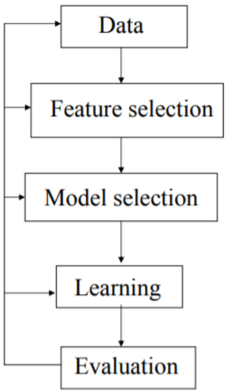
Designing a machine learning system



Model Selection

- selecting the best model from among a collection of candidate machine learning models for the task at hand
- need to select a model that generalizes well on unseen test data points

Designing a machine learning system



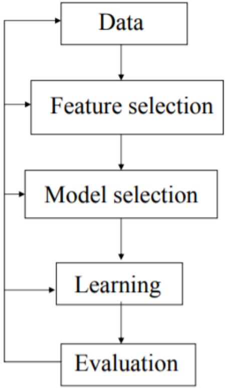
Model Selection

- selecting the best model from among a collection of candidate machine learning models for the task at hand
- need to select a model that generalizes well on unseen test data points

Learning

- using the data to *train* the model

Designing a machine learning system



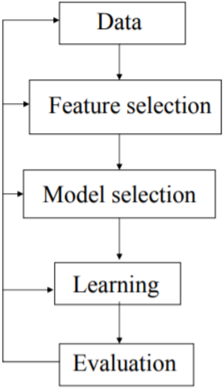
Model Selection

- selecting the best model from among a collection of candidate machine learning models for the task at hand
- need to select a model that generalizes well on unseen test data points

Learning

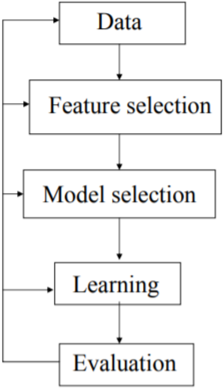
- using the data to *train* the model
- in other words, learning the model parameters from the dataset

Designing a machine learning system



Evaluation

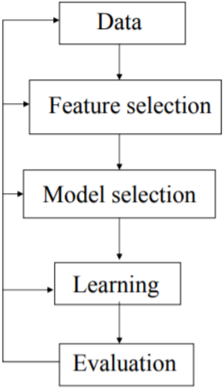
Designing a machine learning system



Evaluation

- measure the performance of the model

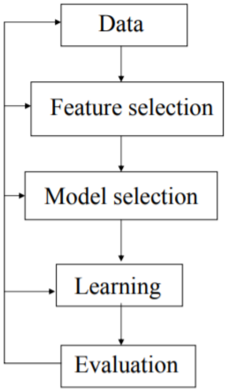
Designing a machine learning system



Evaluation

- measure the performance of the model
- models can be evaluated on multiple metrics
  - eg:- accuracy, precision, etc.

Designing a machine learning system



Evaluation

- measure the performance of the model
- models can be evaluated on multiple metrics
  - eg:- accuracy, precision, etc.
- right choice of an evaluation metric is crucial and often depends upon the problem being solved

References

1. https://en.wikipedia.org/wiki/Artificial_intelligence
2. https://people.cs.pitt.edu/~milos/courses/cs2750-Spring03/lectures/class2.pdf
3. https://christophm.github.io/interpretable-ml-book/terminology.html
4. https://machinelearningmastery.com/neural-networks-are-function-approximators/
5. https://developer.ibm.com/technologies/artificial-intelligence/articles/cc-models-machine-learning/

Thanks Google for the pictures!