

Examples

28 June 2024 11:31

Examples

An **example** is a particular instance of data, (a vector.) We break examples into two categories:

- labeled examples
- unlabeled examples

A **labeled example** includes both feature(s) and label(s).

labeled examples: {features, label}: (x, y)

Use labeled examples to **train** the model. In the case of spam filtering, examples would be individual emails that are labeled "spam."

For example, the following table shows 5 labeled examples of housing information about housing prices in California.

housingMedianAge (feature)	totalRooms (feature)	totalBedrooms (feature)	medianValue (label)
15	5612	1283	645000
19	7650	1901	812000
17	720	174	856000
14	1501	337	714000
20	1454	326	615000

An **unlabeled example** contains features but no label.

x. (We put **x** in boldface to indicate that it is
gories:

and the label. That is:

our spam detector example, the labeled
users have explicitly marked as "spam" or "not

labeled examples from a [data set](#) containing
nia:

medianHouseValue
label)

6900

0100

5700

3400

5500

not the label. That is:

unlabeled examples: {features, ?}: (x, ?)

Here are 3 unlabeled examples from the same dataset, excluding medianHouseValue:

housingMedianAge (feature)	totalRooms (feature)	totalBedrooms (feature)
42	1686	361
34	1226	180
33	1077	271

Once we've trained our model with labeled examples, we can use it to predict a label on unlabeled examples. In the spam dataset, this is useful for identifying spam that humans haven't yet labeled.

me housing dataset, which

examples, we use that model to predict the
detector, unlabeled examples are new emails