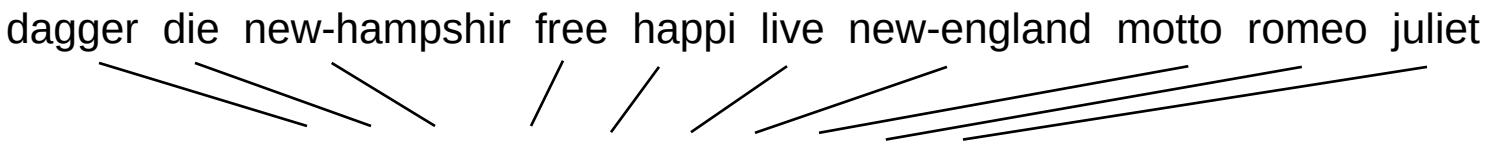
Example: Constructing the Document Matrix

d1 : Romeo and Juliet.
d2 : Juliet: O happy dagger!
d3 : Romeo died by dagger.
d4 : “Live free or die”, that’s the New-Hampshire’s motto.
d5 : Did you know, New-Hampshire is in New-England.



d1	[0, 0, 0, 0, 0, 0, 0, 0, 1, 1]
d2	[1, 0, 0, 0, 1, 0, 0, 0, 0, 1]
d3	[1, 1, 0, 0, 0, 0, 0, 0, 1, 0]
d4	[0, 1, 1, 1, 0, 1, 0, 1, 0, 0]
d5	[0, 0, 1, 0, 0, 0, 1, 0, 0, 0]

Example: Document Classification

- Supervised Learning Problem
  - A human assigns a topic label to each document in a corpus
  - The algorithm learns how to predict the label
- Unsupervised Learning Problem
  - No labels are given
  - Discover groups of similar documents

Learning = Three Core Components

- Representation
- Evaluation
- Optimization

Pedro Domingos, A Few Useful Things to Know about Machine Learning, CACM 55(10), 2012

Learning = Three Core Components

- Representation
  - What exactly is your classifier?
    - A hyperplane that separates the two classes?
    - A decision tree?
    - A neural network?
- Evaluation
- Optimization

Pedro Domingos, A Few Useful Things to Know about Machine Learning, CACM 55(10), 2012

Learning = Three Core Components

- Representation
- Evaluation
  - How do we know if a given classifier is good or bad?
    - # of errors on some test set?
    - Precision and recall?
    - Squared error?
    - Likelihood?
- Optimization

Pedro Domingos, A Few Useful Things to Know about Machine Learning, CACM 55(10), 2012

Learning = Three Core Components

- Representation
- Evaluation
- Optimization
  - How do you search among all the alternatives?
    - Greedy search?
    - Gradient descent?

Pedro Domingos, A Few Useful Things to Know about Machine Learning, CACM 55(10), 2012