SEARCH

RESOURCES

CONCEPTS

1. Video: Introduction
2. Video: Lesson Topics
3. Text: Lesson Topics
4. Video: Latent Features
5. Quiz: Latent Features
6. Video: How to Reduce Features?
7. Video: Dimensionality Reduction
8. Video: PCA Properties
9. Quiz: How Does PCA Work?
10. Screencast: PCA
11. Notebook: PCA - Your Turn
12. Screencast: PCA Solution
13. Screencast: Interpret PCA Results
14. Notebook: Interpretation
15. Screencast: Interpretation Solution
16. Text: What Are EigenValues & Eig...
17. Video: When to Use PCA?
18. Video: Recap
19. Notebook: Mini-Project
20. Mini-Project Solution
21. Video: Outro
22. Text: Recap

Reducing the Number of Features - Dimensionality Reduction

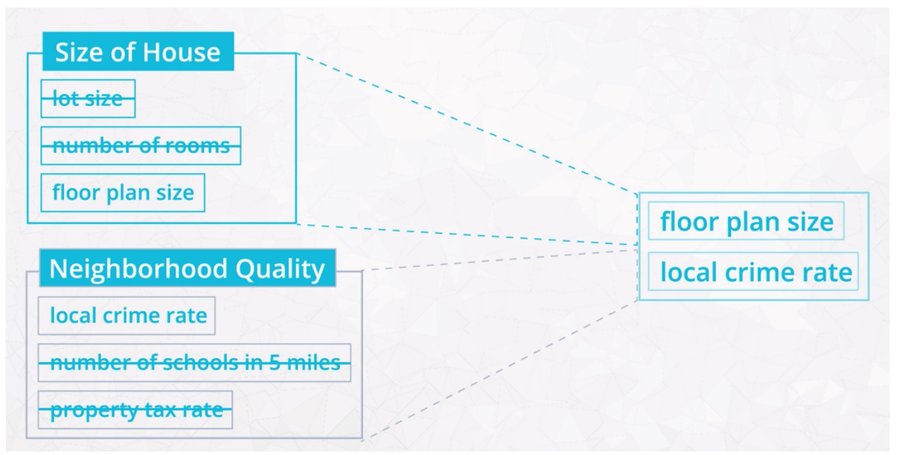
Our real estate example is great to help develop an understanding of feature reduction and latent features. But we have a smallish number of features in this example, so it's not clear why it's so necessary to reduce the number of features. And in this case, it wouldn't actually be required - we could handle all six original features to create a model.

But the **"curse of dimensionality"** becomes more clear when we're grappling with large real-world datasets that might involve hundreds or thousands of features, and to effectively develop a model really requires us to reduce our number of dimensions.

Two Approaches: Feature Selection and Feature Extraction

Feature Selection

Feature Selection involves finding a **subset** of the original features of your data that you determine are most relevant and useful. In the example image below, taken from the video, notice that "floor plan size" and "local crime rate" are features that we have selected as a subset of the original data.



Example of Feature Selection

Methods of Feature Selection:

- **Filter methods** - Filtering approaches use a ranking or sorting algorithm to filter out those features that have less usefulness. Filter methods are based on discerning some inherent correlations among the feature data in unsupervised learning, or on correlations with the output variable in supervised settings. Filter methods are usually applied as a preprocessing step. Common tools for determining correlations in filter methods include **Pearson's Correlation**, **Linear Discriminant Analysis (LDA)**, and **Analysis of Variance (ANOVA)**.
- **Wrapper methods** - Wrapper approaches generally select features by directly testing their impact on the performance of a model. The idea is to "wrap" this procedure around your algorithm, repeatedly calling the algorithm using different subsets of features, and measuring the performance of each model. Cross-validation is used across these multiple tests. The features that produce the best models are selected. Clearly, this is a computationally expensive approach for finding the best performing subset of features, since they have to make a number of calls to the learning algorithm. Common examples of wrapper methods are **Forward Search**, **Backward Search**, and **Recursive Feature Elimination**.

Scikit-learn has a **feature selection module** that offers a variety of methods to improve model accuracy scores or to boost their performance on very high-dimensional datasets.

Feature Extraction

Feature Extraction involves extracting, or constructing, new features called **latent features**. In the example image below, taken from the video, "Size Feature" and "Neighborhood Quality Feature" are new latent features, extracted from the original input data.

Example of Feature Extraction

Methods of Feature Extraction

Constructing latent features is exactly the goal of **Principal Component Analysis (PCA)**, which we'll explore throughout the rest of this lesson.

Other methods for accomplishing Feature Extraction include **Independent Component Analysis (ICA)** and **Random Projection**, which we will study in the following lesson.

Further Exploration

If you're interested in a deeper study of these topics, here are a couple of helpful blog posts and a research paper:

- https://www.analyticsvidhya.com/blog/2016/12/introduction-to-feature-selection-methods-with-an-example-or-how-to-select-the-right-variables/
- https://elitedatascience.com/dimensionality-reduction-algorithms
- http://www.ai.mit.edu/projects/jmlr/papers/volume3/guyon03a/source/old/guyon03a.pdf