

Feature Engineering

Papangkorn Inkeaw, Ph.D.

Feature Construction

Chapter 4 (Part I)

Feature Construction

- Manually creating new features by
 - Transforming existing features
 - Discretization
 - Categorical variable encoding
 - Joining some features together
 - Feature combining using statistical operations
 - Feature combining using mathematical functions
 - Feature combining using decision trees
 - Expand a feature
 - Polynomial Expansion
- Feature construction is often laborious and time-consuming
- It demands the most domain knowledge.
 - Impossible to hand-craft features without a deep understanding of the underlying problem domain.

Feature Combining

- Variables can be combined by using domain knowledge of the data and the industry.
- For example, in finance, we combine information about the income and the acquired debt to determine the disposable income:

$$\text{disposable income} = \text{income} - \text{total debt}$$

- Similarly, total debt can be derived by adding all debts up:

$$\text{total debt} = \text{car loan balance} + \text{credit card balance} + \text{mortgage balance}$$

Feature Combining using Statistical Operations

- New features can be created by performing statistical operations over existing variables.
 - Mean, median and mode
 - Standard deviation and variance
 - Skewness and kurtosis
 - Maximum and minimum
 - Percentile
- The variables that are combined should be the same things or comparable but differences in times, positions or aspects.

Feature Combining using Statistical Operations

Example: Weather

The temperature can be collected from an IoT device every 30 minutes. For a day, thus, we have 48 temperature values. However, using all the 48 temperature values as predictors may not be suitable.

- We can average the temperature values to capture their trend.
- The standard deviation or variance can be used to explain their dispersion.
- Maximum or minimum values can be used when we emphasize the extreme value.

Feature Combining using Statistical Operations

Example: Medicine

The decision for surgical intervention for craniotomy is dependent on the injury type and the patient's neurologic exam. The volume and thickness of hemorrhage subtypes are used as indicators. For a patient, hemorrhage may present in many regions. So, we have many volumes and thicknesses for each hemorrhage region.

- Maximum volume and thickness for each hemorrhage subtype are used as indicators.

Feature Combining using Statistical Operations

Example: Website analysis

The time-on-page for a web page is the time difference between the point when a person lands on the page and when they move on to the next one. As each website contains more than web pages, it has many time-on-page values. We want to analyze the users' attention on some websites.

- We can sum up all time-on-page of all pages to make a feature.
- Minimum or maximum time-on-page may be useful information.

Combining Pairs of Features with Mathematical Functions

- Some mathematical operations, make more sense when performed between two features, or when considering multiple features against one reference variable.
 - Subtraction explains the difference between two features
 - Perform the absolute function when considering only magnitude
 - Division explains the ratio of a feature to a reference feature
- For example, debt to income ratio:

$$\text{debt to income ratio} = \frac{\text{total debt}}{\text{total income}}$$

Combining Pairs of Features with Mathematical Functions

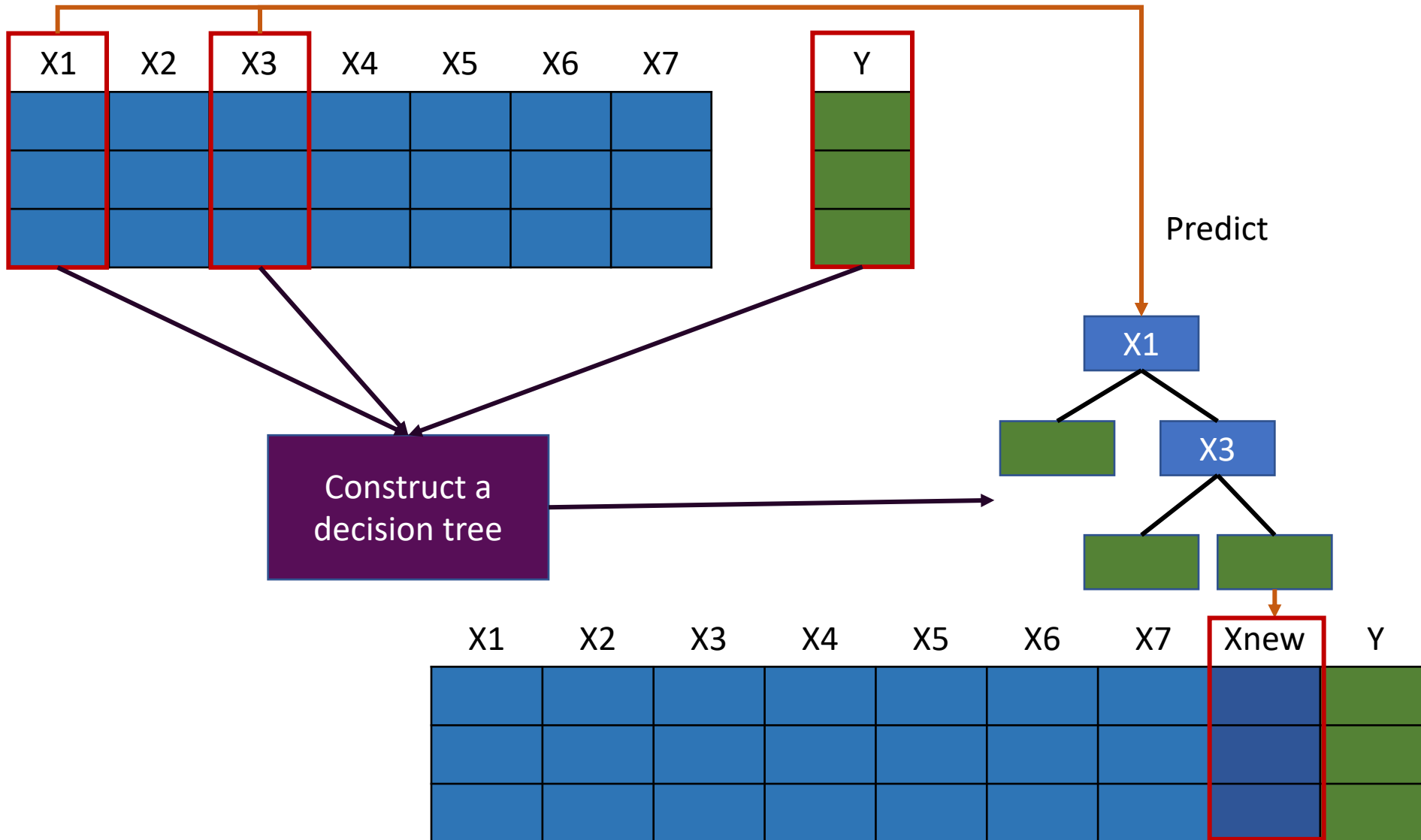
- For example, disposable income

$$\textit{disposable income} = \textit{income} - \textit{total debt}$$

Feature Combining using Decision Trees

- Alexandru Niculescu-Mizil, et. al., proposed a method for combining two or more variables using decision trees. (Wined the KDD competition in 2009)
- The procedure consists of
 - Building a decision tree using a subset of the features, (two or three at a time)
 - Using the prediction of the tree as a new feature.
- It creates new features with monotonic relationships between features and target.
- It also captures feature interactions.

Feature Combining using Decision Trees



References & Study Resources

- Soledad Galli. (2020). *Python Feature Engineering Cookbook*. Packt Publishing.
- Inkeaw, P., Angkurawaranon, S., Khumrin, P., Inmutto, N., Traisathit, P., Chaijaruwanich, J., Angkurawaranon, C., & Chitapanarux, I. (2022). Automatic hemorrhage segmentation on head CT scan for traumatic brain injury using 3D deep learning model. *Computers in biology and medicine*, 146, 105530.
- <http://www.mtome.com/Publications/CiML/CiML-v3-book.pdf>
- <https://www.globalmediainsight.com/blog/avg-time-on-page-vs-session-duration/>