Regression

Exercise 6
Recap: Regression

- Classification predicts a *nominal* value
  - A finite set of values
- Regression predicts a *numerical* value
  - A possibly infinite set of possible values
  - Can be *interpolating* and *extrapolating*

[Diagram: My Hobby: Extrapolating]

http://xkcd.com/605/
K Nearest Neighbours Regression

• Find the k nearest neighbours
• And use the average of their label as prediction
• Only interpolating regression possible

• It’s the same operator that you already know from classification!
Regression Trees / SVM / ANN

• Other operators that you already know and which can be used for regression:

• Decision Tree
  • Set criterion to „least square“

• SVM
  • Set svm type to „epsilon-SVR“ or „nu-SVR“

• Neural Net
  • No changes required
Linear Regression

• Finds a linear function
  \[ f(x) = w_0 + w_1 x_1 + w_2 x_2 + \ldots + w_n x_n \]

• That minimises the error
  \[ \sum_{\text{all examples}} \left( w_0 + w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_n \cdot x_n - y \right)^2 \]
Operators: Linear Regression

• A learning operator that learns a linear regression model
  • Selects the features automatically

• Parameters:
  • Feature selection:
    • none, M5 prime, greedy, T-Test, Iterative T-Test
  • Use bias:
    • determines if an intercept should be used in the regression
  • Ridge:
    • Controls the slope of the learned function, higher ridge results in smaller coefficients

\[
\sum_{\text{all examples}} \left( w_0 + w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_n \cdot x_n - y \right)^2 + \lambda \sum_{\text{all variables}} w_i^2
\]
Operators: Polynomial Regression

Original attributes are transformed before running a linear regression

- **Parameters:**
  - **Replication factor:**
    - How often can a feature be replicated in the transformation?
  - **Max degree:**
    - Maximal degree of the final polynomial
  - **Min/Max coefficient:**
    - Limit the values of the coefficients

\[ y(x, w) = w_0 + w_1 x + w_2 x^2 + \ldots + w_M x^M \]
Operators: Polynomial Regression

• Careful: the polynomial regression operator does not always produce the expected result!

• Alternative: Manually create a polynomial regression
  • Using the generate attributes operator
  • And a linear regression afterwards
Operators: Local Regression

- Lazy Learning!
- Retrieves the k nearest neighbours, calculates a regression model, then predicts the value
- Parameters:
  - Degree:
    - Degree of the locally fitted polynomial
  - Measure, neighbourhood type, k:
    - Used to select the nearest neighbours
Performance Measures for Regression

• **Mean Absolute Error**
  • How far are we off on average?
  $$\text{MAE} = \frac{\sum_{\text{all examples}} |\text{predicted} - \text{actual}|}{N}$$

• **Root Mean Squared Error**
  • Re-scales the errors:
    • Large errors have more influence
    • Small errors have less influence
  $$\text{RMSE} = \sqrt{\frac{\sum_{\text{all examples}} (\text{predicted} - \text{actual})^2}{N}}$$

• **Coefficient of Determination (R²)**
  • Tells you how much of the variation of your target variable is explained by the model
  $$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$