Module 3

Linear and Nonlinear Regression

▼ What does regression predict?

numerical values

▼ Regression illustrates a relationship between what?

a set of static independent variables and a dependent variable

▼ What do regression algorithms estimate?

the mapping function from the input variable to continuous output variables (usually numerical)

▼ What is linear regression?

- In linear regression, a linear relationship is modeled between a dependent (scalar) variable and one or more independent (explanatory) variables (also known as predictors).
- The dependent variable is the feature being predicted (the outcome variable).
- If there is only one independent variable, it is a ***simple linear regression*** model.
- Linear regression holds a high danger of overfitting.
- Note: most appropriate for when outcome variable is **continuous**

▼ What is logistic regression?

Logistic regression is actually a form of binary classification that uses the same underlying technique used in linear regression, with the difference being that the output of logistic regression is a probability reducible to a discrete binary.

▼ What is polynomial regression?

Polynomial regression finds the relationship between x and the conditional mean of the corresponding value of y.

▼ When is polynomial regression effective?

- Polynomial Regression is effective when your data points have a curvilinear relationship, meaning that as **one of two related variables increases the other does as well, but only up to a point**, at which the distance between the variables increases as they diverge.
- Curvilinear relationships can be quadratic or cubic.
- If one variable can be squared to account for variability in the other, you will see a curvilinear effect.

▼ What is the formula for linear regression?

Y=bX+m, where the intercept is *b*, and the line slope is *m*

▼ What is the ordinary least squares (OLS) method?

a method of fitting the decision boundary in linear regression

▼ How does OLS work?

It finds the optimum hyperplane that minimizes the sum of the squared *residual*, or the distances between data points and the hyperplane.

The expected conditional mean value can then be estimated.

▼ What is one way of preventing overfitting?

the use of regularization methods

▼ What is early stopping?

where you monitor the generalization error and ending the training process when that error rate begins to degrade

▼ What are some techniques for finding optimum hyperplanes in your data using linear regression?

- Random guess: picking the best possibility after a number of random guesses.
- Random step: it moves the hyperplane up or down a step randomly, checking the line each time.
- Smart step: your model will try a random hyperplane, then use gradient descent methods to limit the next step in the search.
- Calculated shortcut: uses mathematical tricks to guess the best hyperplane, for instance by finding the mean of the target distance, or by using the least squares fitting method.

▼ How can K-NN techniques can be used for regression?

Instead of voting for a possible answer based on distance, you instead combine the values the outputs represent into summary statistics such as the median or the mean.

▼ What is stepwise regression?

Stepwise regression is a relatively simple technique that uses forward selection, backward elimination, or bidirectional elimination to add or subtract from the set of explanatory (independent) variables each step.

▼ What is ridge regression?

- a type of multiple logistic regression model
- bias is added to your regression estimates
- used when multicollinearity occurs (two or more predictor variables are highly correlated)

▼ What is discriminant analysis?

When multiple variables must be taken into account, discriminant analysis (remember Fisher's rule?) is used to find the variables that discriminate one (or more) group(s) from another.

▼ What is Lasso regression?

A form of regression that selects only a portion of the provided independent variables, and provides regularization by using *soft thresholding.*

▼ What can Bayesian inference be used for?

- To estimate the posterior probability distributions of the models variables.
- This is most commonly used when errors of the model have a normal distribution.
- This can be expressed as posterior = (Likelihood * Prior)/Normalization.
- The model essentially becomes non-parametric and domain knowledge should be used.