**ML revision notes**

1. Linear Regression:

Linear regression models the relationship between a dependent variable y and one or more independent variables “x\_1, x\_2, ..., x\_n” by fitting a linear equation to the observed data.

- Formula: y = b0 + b1x1 + b2x2 +... + e

- Assumptions:

- Linearity: y is linearly related to the predictors.

- Independence of Observations: Observations are independent.

- Homoscedasticity: Constant variance of errors.

- Limitations:

- Linear Assumption.

- Sensitivity to Outliers.

- Multicollinearity.

2. Multiple Regression:

- \*\*Definition:\*\* Multiple regression extends linear regression to model the relationship between a dependent variable and multiple predictors.

- \*\*Formula: y = B0 + B1 x1 + B2 x2 + ... + \Bn xn + e

- Assumptions:Similar to linear regression but extended to multiple predictors.

- Limitations: Same as linear regression, but also includes challenges related to interpreting coefficients with multiple predictors.

3. Logistic Regression:

- Definition: Logistic regression predicts the probability of a binary outcome based on one or more predictors.

- Formula: ( P(y=1|x) = \frac{1}{1 + e^{-(B0 + B1 x\_1 + B2 x\_2 + ... + Bn xn)

4. Naive Bayes:

Naive Bayes is a probabilistic classifier based on Bayes' theorem and the assumption of conditional independence between features.

- Formula:( P(y|x) = frac{P(x|y)P(y)}{P(x)}

-Assumptions: Assumes features are conditionally independent given the class label.

- Limitations:

- Overly Simplistic Assumption.

- Sensitivity to Irrelevant Features.

5. Support Vector Machines (SVM):

- Definition: SVM finds the optimal hyperplane to separate classes with maximum margin.

- Assumptions:\*\* Aims to maximize the margin between classes.

kernels in SVM are:

Linear Kernel: Suitable for linearly separable data or when the number of features is large compared to the number of samples. It works well for simple classification tasks with linear decision boundaries.

Polynomial Kernel: Useful for capturing non-linear relationships between features. It introduces polynomial terms to the decision function, allowing SVM to fit more complex decision boundaries.

Radial Basis Function (RBF) Kernel: Most commonly used kernel for SVM. It is versatile and effective in capturing non-linear relationships. RBF kernel maps data into a high-dimensional space using Gaussian radial basis functions, enabling SVM to find complex decision boundaries.

Sigmoid Kernel: Suitable for binary classification problems where the relationship between features is similar to a sigmoid function. It can be useful for neural network-like architectures but is less commonly used compared to linear, polynomial, and RBF kernels.

Margin: The margin is the distance between the decision boundary (hyperplane) and the closest data points from each class, known as support vectors. SVM aims to maximize this margin to improve generalization to unseen data.

Support Vectors: Support vectors are the data points that lie closest to the decision boundary and directly influence the position and orientation of the hyperplane. They are crucial for defining the margin and determining the optimal separation between classes.