

P S DS ML AI

May 15, 2023

Contents

1	Supervised Learning	2
1.1	Classification	2
1.1.1	KNN	2
2	Measuring Model Performance	2

Machine Learning

1 Supervised Learning

1.1 Classification

1.1.1 KNN

K-Nearest-Neighbours - The euclidean distance between a data point and a series of features is calculated and the majority value of the target variable is taken.

$$distance = \sqrt{(w_0 - w_1)^2 + (x_0 - x_1)^2 + (y_0 - y_1)^2 + (z_0 - z_1)^2 \dots}$$

- A large value of k leads to a less complex model and can lead to *underfitting*
- A smaller value of k increases the complexity of the model and can lead to *overfitting*

2 Measuring Model Performance