MINDD Formulary

1 Normalization Functions

Min-max Normalization
$$y^{'} = \frac{y - min_y}{max_y - min_y} \tag{1}$$

Zscore Normalization
$$y^{'} = \frac{y - \mu_y}{\sigma_y} \qquad \begin{array}{l}\mu_y \text{ média y} \\ \sigma_y \text{ desvio padrão y}\end{array} \tag{2}$$

Sigmoidal Normalization
$$y^{'} = \frac{1 - e^{-\beta}}{1 + e^{-\beta}}, \qquad \beta = \frac{1 - \mu_y}{\sigma_y} \tag{3}$$

2 Distance Functions

Eucledian Distance
$$dist(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2} \tag{4}$$

Minkowski Distance
$$dist(x, y) = \sqrt[m]{|x_1 - y_1|^m + |x_2 - y_2|^m + ... + |x_n - y_n|^m} \tag{5}$$

Manhattan Distance
$$dist(x, y) = |x_1 - y_1| + |x_2 - y_2| + ... + |x_n - y_n| \tag{6}$$

Hamming Distance
$$dist(x, y) = \sum_{i=1}^{n} x_i \neq y_i \tag{7}$$

3 Conditional Probability - Bayes Theorem

Bayes Theorem
$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)} \tag{8}$$

Naive Bayesian Classifier
$$V_{MPP} = \arg max_{v_j \in V} P(v_j) \times \prod P(A_i|v_j) \tag{9}$$

Continuous Attributes
$$P(A_i|c_j) = \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} e^{\frac{-(A_i - \mu ij)^2}{2\sigma_{ij}^2}} \tag{10}$$

4 Model Evaluation

Accuracy
$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

Error Rate
$$ErrorRate = 1 - accuracy \tag{12}$$

Sensitivity
$$TPR = \frac{TP}{TP + FN} \tag{13}$$

Specificity
$$TNR = \frac{TN}{TN + FP} \tag{14}$$

False Positives Rate (1 - Specificity)
$$FPR(fallout) = \frac{FP}{FP + TN} \tag{15}$$

Precision
$$precision = \frac{TP}{TP + FP} \tag{16}$$

Recall
$$recall = \frac{TP}{TP + FN} \tag{17}$$

F1
$$F1 = \frac{2 \times precision \times recall}{precision + recall} = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{18}$$

5 Association Rules

Support
$$Sup(A \Rightarrow B) = \frac{P(A \cap B)}{n} \tag{19}$$

Confidence
$$Conf(A \Rightarrow B) = \frac{P(A \cap B)}{P(A)} \tag{20}$$

Coverage
$$Coverage(A \Rightarrow B) = \frac{Sup(A \Rightarrow B)}{Conf(A \Rightarrow B)} \tag{21}$$

Interest
$$Interest(A \Rightarrow B) = \frac{Sup(A \Rightarrow B)}{Sup(A) \times Sup(B)} \tag{22}$$

Leverage
$$Leverage(A \Rightarrow B) = Sup(A \Rightarrow B) - Sup(A) \times Sup(B) \tag{23}$$

Convinction
$$Convinction(A \Rightarrow B) = \frac{1 - Sup(B)}{1 - Conf(A \Rightarrow B)} \tag{24}$$

6 Text Mining

TF-IDF
$$TF - IDF = \frac{N_t}{N_d} \times log_{10}(\frac{N_{dt}}{N_{dc}}) \tag{25}$$