**Decision Tree Algorithm**

Decision tree algorithm is a data mining induction techniques that recursively partitions a data set of records using depth-first greedy approach (Hunts et al, 1966) or breadth-first approach (Shafer et al, 1996) until all the data items belong to a particular class. A decision tree structure is made of root, internal and leaf nodes. The tree structure is used in classifying unknown data records. At each internal node of the tree, a decision of best split is made using impurity measures (Quinlan, 1993). It is used to improve the prediction and classification accuracy of the algorithm by minimizing over-fitting (noise or much detail in the training data set) (Mehta et al, 1996). Decision tree classifiers obtain similar and sometimes better accuracy when compared with other classification methods (1) (2).

Iterative Dichotomiser 3 is a simple decision tree learning algorithm introduced in 1986 by Quinlan Ross. It is serially implemented and based on Hunt´s algorithm. The basic idea of ID3 algorithm is to construct the decision tree by employing a top-down, greedy search through the given sets to test each attribute at every tree node.

**Linear Discriminant Analysis (LDA)**

LDA would separate the two normally distributed classes well. This method used to find a linear combination of features that characterizes or separates classes. In LDA, the dimensional embeddings are reduced in such a way that the orientation of the projected data of classes on an arbitrary line or space is well-separated from each other. The resulting combination is used for dimensionality reduction before classification (3) (4).

**Naive Bayes**

The Naive Bayes algorithm is an intuitive method that uses the probabilities of each attribute belonging to each class to make a prediction based on Bayes Theorem. It is the supervised learning approach you would come up with if you wanted to model a predictive modeling problem probabilistically. Naive bayes simplifies the calculation of probabilities by assuming that the probability of each attribute belonging to a given class value is independent of all other attributes. The algorithm allows us to predict a class, given a set of features using probability. This is a strong assumption but results in a fast and effective method (5).

# Bibliografía

1. *Data Mining with Decision Trees and Decision Rules.* **Apté, Chidanand y Weiss, Sholom M.**

2. *A Heuristic-Based Decision Induction Method for Noisy Data.* **Kerdprasop, Nittaya y Kerdprasop, Kittisak.** Thailand : Communications in Computer and Information Science, 2011.

3. **Haeb-Umbach, R. y Ney , H.** *Linear Discriminant Analysis for improved large vocabulary continuous speech recognition.* Aachen : Philips Research Laboratory Aachen.

4. *Class-dependent PCA, MDC and LDA: A combined classifier for pattern classification.* **Sharma, Alok, Paliwal, Kuldip K y Onwubolu, Godfrey C.** 7, s.l. : ELSEVIER, 2006, Vol. 39. ISSN 0031-3203.

5. *Learning accurate and concise naïve Bayes classifiers from attribute value taxonomies and data.* **Zhang, J., Kang, D.-K., Silvescu, A., & Honavar, V.** s.l. : Knowledge and Information Systems, 2006, Vol. 9(2). http://doi.org/10.1007/s10115-005-0211-z.

6. *Linear discriminant analysis for improved large vocabulary continuous speech recognition. In Acoustics, Speech, and Signal Processing.* **Ney, H. y Haeb-Umbach, R.** ICASSP-92., s.l. : IEEE International Conference, 1992, Vols. Vol. 1, pp. 13-16.