

ABSTRACT

ANDRADE, Domingos Sávio Mattos de. *Detection of respiratory changes in systemic sclerosis by forced oscillation technique and machine learning algorithms*. 119f. 2018. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2018.

The Forced Oscillation Technique (TOF) has shown high potential in the study of mechanical alterations of respiratory diseases. Due to its simplicity of execution, requiring little patient's collaboration, it provides new parameters, which complement the pulmonary evaluations performed by traditional methods, such as spirometry. To prove the efficiency of this new technique, this work proposes the use of machine learning algorithms help the investigation and the early diagnosis of respiratory changes in patients with Systemic Sclerosis. The data provided by TOF were used on the following algorithms: K-Nearest Neighbor (KNN), Radial Support Vector Machine (RSVM), Adaboost (ADAB), Random Forest (RF) and Multiple Instance Learning (AMI) as classifier and attribute selector. From the experiments performed, dynamic compliance (C_{din}) was the attribute that presented the best individual performance (AUC = 0,77 for the control group x normal spirometric examination (group 1) and 0,94 for the control group x restrictive (group 2)), indicating moderate and high accuracy, respectively. In the experiment with all TOF attributes, the best performance was achieved by KNN (AUC = 0,90) for group 1, indicating high accuracy and AUC = 0,96, achieved by ADAB for group 2. The AMI algorithm obtained AUC = 0,72 for group 1 and AUC = 0,94 for group 2. With five attributes selected by the multi-instance algorithm, the best result was achieved by the KNN algorithm (AUC = 0,88) for group 1 and ADAB (AUC = 0,98) for group 2. As a classifier, the AMI obtained AUC of 0,74 and 0,95, respectively. In the experiment performed with five attributes selected by the expert, the best performance was achieved by KNN (AUC = 0,85) for group 1 and ADAB (AUC = 0,98) for group 2. With direct selection, the best performance was also achieved by KNN for group 1 (AUC = 0,89) and by ADAB for group 2 (AUC = 0,97). The experiments showed that the use of machine learning algorithms increased the accuracy of the diagnosis of Systemic Sclerosis, which may help in the early diagnosis of the disease.

Keywords: Systemic Sclerosis; Forced Oscillations Technique; TOF; Machine learning; Multiple Instances; AUC.