

# 1 Resumo

Este trabalho tem como objetivo principal adaptar o modelo STR-Tree, o qual é a combinação de um modelo *Smooth Transition Regression* com *Classification and Regression Tree (CART)*, a fim de utilizá-lo em Classificação. Para isto algumas alterações foram realizadas em sua forma estrutural e na estimação. Devido ao fato de estarmos fazendo classificação de variáveis dependentes binárias, se faz necessária a utilização das técnicas empregadas em Regressão Logística, dessa forma a estimação dos parâmetros da parte linear passa a ser feita por Máxima Verossimilhança. Assim o modelo, que é paramétrico não-linear e estruturado por árvore de decisão, onde cada nó terminal representa um regime os quais têm seus parâmetros estimados da mesma forma que em uma Regressão Logística, é denominado *Smooth Transition Logistic Regression-Tree (STLR-Tree)*. A inclusão dos regimes, determinada pela divisão dos nós da árvore, é feita baseada em testes do tipo Multiplicadores de Lagrange, que em sua forma para o caso Gaussiano utiliza a Soma dos Quadrados dos Resíduos em suas estatísticas de teste, aqui são substituídas pela Função Desvio (*Deviance*), que é equivalente para o caso dos modelos não Gaussianos, cuja distribuição da variável dependente pertença à família exponencial. Na aplicação a dados reais selecionou-se dois conjuntos das variáveis explicativas de cada uma das duas bases utilizadas, que resultaram nas melhores taxas de acerto, verificadas através de Tabelas de Classificação (Matrizes de Confusão). Esses conjuntos de variáveis foram usados com outros métodos de classificação existentes, são eles: *Generalized Additive Models (GAM)*, *Regressão Logística*, *Redes Neurais*, *Análise Discriminante*, *k-Nearest Neighbor (k-NN)* e *Classification and Regression Trees (CART)*. O STLR-Tree apresentou o segundo melhor resultado para a taxa de acerto total em três das quatro aplicações mostradas, ficando, em todas elas, atrás apenas de GAM. Além disso, as estruturas das árvores foram pequenas, não passando de dois nós terminais, indicando assim que o modelo, além dos bons resultados frente aos outros métodos, foi parcimonioso.