

Aplicação de Modelos Mistos e SVM em Seleção Genômica de árvores de Eucalipto

Ana Gabriela Pereira de Vasconcelos ^A, Joanlise Marco Leon de Andrade ^B

^A Instituto de Matemática e Estatística – Universidade de São Paulo

^B Departamento de Estatística – Universidade de Brasília

Programas de melhoramento genético de árvores de floresta visam aumentar a qualidade e ganho econômico de suas plantações por meio de manipulação genética, porém essa tarefa envolve desafios como longos ciclos de cruzamento e altos custos de coleta de fenótipos. Nesse sentido, abordagens que avaliam valores genéticos de árvores jovens, sem a necessidade de fenotipagem, possuem o potencial de superar estes desafios. Uma delas é a Seleção Genômica, que consiste em utilizar informações moleculares para estimar efeitos de marcadores genéticos com base em um modelo de predição. O modelo, desenvolvido em uma população de treinamento com informações genotípicas e fenotípicas, é utilizado para obter valores genéticos baseados em dados genotípicos de plantas candidatas, portanto, a escolha do modelo é uma etapa essencial. Este estudo compara modelos mistos e SVMs em dados de eucaliptos, além de estudar fatores que influenciam as métricas obtidas, como características genéticas, qualidade dos fenótipos e efeitos de parentesco. Notou-se que os modelos para os fenótipos com maiores herdabilidades apresentaram medidas de previsão também superiores. Ainda foi possível verificar a importância de controlar os efeitos de parentesco por meio da validação cruzada para a obtenção de métricas menos otimistas, uma vez que os modelos serão utilizados com dados de novos indivíduos que não estavam presentes na população de treinamento. Por fim, observou-se que os modelos de regressão e de SVM apresentaram resultados consistentes, os quais evidenciaram que sua escolha deve depender do estudo em questão.

Palavras-chave: Seleção genômica, eucalipto, melhoramento genético, machine learning, regressão ridge, validação cruzada, SVM.

Notou-se que os modelos para os fenótipos com maiores herdabilidades apresentaram medidas de previsão também superiores.