Web-based analysis of genotypic, phenotypic, and environmental data for forest trees

Forest trees are long-lived and immobile individuals that serve as ideal models to assess population structure and adaptation to the environment. Despite the availability of comprehensive data, integrating the data describing genotype, phenotype, and the environment is still challenging. To address this, CartograTree was designed and implemented as an open repository and open-source analytic web-based framework for all three.

The source code is available on GitLab, and the documentation on Read the Docs. This project is currently funded through a USDA NIFA FACT grant.
Machine learning algorithms for domain adaptation

These algorithms are designed to address the lack of labeled data in a new domain, which makes supervised machine learning algorithms virtually useless. They do so by integrating data from a related domain, that has plenty of labeled instances. We evaluated these algorithms with biological sequences for tasks of splice site prediction, and protein localization, as well as with Twitter data for disaster management.