

FOREST COVERTYPE CLASSIFICATION

MATEJ KVAŠŠAY

PROJECT DESCRIPTION

PROJECT GOALS

- ▶ Train various models for classifying forest cover into 7 classes
- ▶ Evaluate performance, compare models and investigate which attributes contribute the most
- ▶ Demonstrate usage of methods introduced in this course

DATASET

- ▶ "Coverture data set" from UCI ML repository (1998)
- ▶ Over 580000 data instances
- ▶ 54 cartographic attributes
- ▶ Multivariate data - quantitative (elevation, slope, aspect, hillshade) and binary (soil types)

PROPOSED APPROACH

- ▶ Statistical analysis of attribute values
- ▶ Feature selection
- ▶ Training SciKit-Learn models with multiple configurations
- ▶ Cross validation