Completed • Knowledge • 1,694 teams

Forest Cover Type Prediction

Fri 16 May 2014 – Mon 11 May 2015 (2 months ago)

Dashboard

Home
- Data
- Make a submission

Information
- Description
- Evaluation
- Rules

Forum

Scripts

Leaderboard

My Team
- GitHub

My Submissions

Leaderboard

1. antgleb
2. Ashish Singh
3. ucbw207_2_forest
4. Tian Zhou
5. Doru Arfire
6. TIEN VU
7. Eugene Gritskevich
8. Ivan Liu
9. Céline Theeuws
10. Deep Learner (Rahul Mohan)

53 Scripts

First try with Random Forests
0 Votes / 7 days ago / Python

First try with Random Forests
2 Votes / 34 days ago / Python

testing
0 Votes / 2 months ago / Python

testing d3.js
0 Votes / 2 months ago / Python

RandomForest
0 Votes / 2 months ago / Python

RMarkdown Starter Script
5 Votes / 2 months ago / RMarkdown

Competition Details » Get the Data » Make a submission

Use cartographic variables to classify forest categories

Get started on this competition with Kaggle Scripts. No data download or local environment needed!

Random forests? Cover trees? Not so fast, computer nerds. We're talking about the real thing.

In this competition you are asked to predict the forest cover type (the predominant kind of tree cover) from strictly cartographic variables (as opposed to remotely sensed data). The actual forest cover type for a given 30 x 30 meter cell was determined from US Forest Service (USFS) Region 2 Resource Information System data. Independent variables were then derived from data obtained from the US Geological Survey and USFS. The data is in raw form (not scaled) and contains binary columns of data for qualitative independent variables such as wilderness areas and soil type.

This study area includes four wilderness areas located in the Roosevelt National Forest of northern Colorado. These areas represent forests with minimal human-caused disturbances, so that existing forest cover types are more a result of ecological processes rather than forest management practices.

Acknowledgements

Kaggle is hosting this competition for the machine learning community to use for fun and practice. This dataset was provided by Jock A. Blackard and Colorado State University. We also thank the UCI machine learning repository for hosting the dataset. If you use the problem in publication, please cite:

Bache, K. & Lichman, M. (2013). UCI Machine Learning Repository. Irvine, CA: University of California, School of Information and Computer Science

Started: 6:55 pm, Friday 16 May 2014 UTC
Ended: 11:59 pm, Monday 11 May 2015 UTC (360 total days)
Points: this competition did not award ranking points
Tiers: this competition did not count towards tiers

Forum (46 topics)

Validation versus LB score
31 days ago

How is evaluation of results calculated?
32 days ago

Final results!
48 days ago

e+05
2 months ago

Forecasting ARIMA help
2 months ago

I'm liking Seaborn
2 months ago

teams

players

entries