Mac Barth
mbarth@regis.edu
(605) 786-3370
MSDS 696 Practicum Proposal

**Detecting Forest Canopy Change Over Time in Colorado**

**Description**: This project investigates changes in forest canopy height over time using multi-year CHM (Canopy Height Model) data from NEON field sites in Colorado. The analysis focuses on identifying areas of significant canopy loss or growth, which may indicate ecological disturbance (e.g., fire, disease, land use) or forest recovery. Exploratory Data Analysis (EDA) will be used to summarize and visualize canopy height differences, while a supervised classification task will categorize areas by dominant canopy change type or land cover class such as forest, shrubland, or grassland.

**Data Science Task**:
- Geospatial data visualization and EDA
- Change detection via raster differencing
- Supervised classification using CHM-derived features

**Data**:
- **Primary dataset**: NEON CHM (Canopy Height Model) data (DP3.30015.001) for Colorado NEON sites
- **Secondary**: NEON vegetation indices (DP3.30006.001) and site event logs (DP1.10111.001)
- **Data volume**: Each CHM tile is ~30–50 MB in memory. Expect to use 4–8 tiles from two timepoints, totaling ~500 MB.
- All data is publicly available from the NEON Data Portal.

**Analysis Method**:
- Perform EDA using Python to load, align, and compare CHM rasters from two years
- Generate difference rasters and spatial plots of canopy height change
- Classify raster patches into categories (growth, loss, stable) using Random Forest or Decision Tree models
- Explore classification into land cover types (forest, shrubland, grassland) using CHM + vegetation indices

**Anticipated Difficulties**:

- **Multi-year CHM availability:** May be limited across sites. Will focus only on Colorado sites that meet multi-year coverage criteria
- **Raster alignment issues:** Addressed through reprojection and masking invalid values
- **Memory limitations:** Will process data in patches and use efficient raster operations

**Timeline**

| | |
|---|---|
| Week 1-2 | Finalize project proposal. Select Colorado NEON site(s) with multi-year CHM coverage. Review NEON documentation and data product structure. |
| Week 3-4 | Download CHM data from NEON for selected years. Load and align raster tiles in Python. Mask invalid data values and prepare raster grids for analysis. |
| Week 4-5 | Perform exploratory data analysis: generate difference rasters (ΔCHM), visualize spatial patterns, and create histograms and summary statistics of canopy change. Identify zones of significant growth or loss. |
| Week 6-7 | Develop and train a supervised classification model (e.g., Random Forest) to categorize raster patches into "growth," "loss," or "stable" classes. Evaluate model accuracy and refine input features (e.g., vegetation indices, texture metrics). |
| Week 8 | Finalize visualizations, documentation, and GitHub repo. Prepare and deliver final project presentation and summary report. |

**Github: [macfredough/co-chm-analysis: Detecting and classifying forest canopy height changes over time in Colorado NEON sites using LiDAR-derived CHM data and machine learning.](#)**