TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT
KHOA CÔNG NGHỆ THÔNG TIN



TRẦN TRUNG CHUYÊN

BÁO CÁO SINH HOẠT HỌC THUẬT

Hà Nội, 12-2021

TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT
KHOA CÔNG NGHỆ THÔNG TIN

BÁO CÁO SINH HOẠT HỌC THUẬT

Đề tài: "Xử lý dữ liệu không gian địa lý, hình ảnh vệ tinh quy mô lớn sử dụng nền tảng đám mây Google Earth Engine"

Người báo cáo: TS. Trần Trung Chuyên
Đơn vị: Bộ môn Tin học trắc địa

Hà Nội, 12-2021

MỤC LỤC

MỤC LỤC **5**

GIỚI THIỆU **7**

CHƯƠNG 1 GIỚI THIỆU NỀN TẢNG ĐÁM MÂY GOOGLE EARTH ENGINE **9**

1.1 TẠO TÀI KHOẢN ĐỂ LÀM VIỆC VỚI GEE 9
1.2 ĐĂNG NHẬP ĐỂ LÀM VIỆC VỚI GEE 10
1.3 GIỚI THIỆU TRÌNH SOẠN THẢO MÃ CỦA GEE 10
1.4 HỌC MÁY TRONG GEE 11

CHƯƠNG 2 MỘT SỐ CHỨC NĂNG CƠ BẢN CỦA GOOGLE EARTH ENGINE **13**

2.1 CĂN GIỮA KHU VỰC ĐÃ BIẾT TỌA ĐỘ VÀ MỨC THU PHÓNG 13
2.2 THÊM LỚP DỮ LIỆU ẢNH SENTINEL-2 14
2.3 NHẬP RANH GIỚI KHU VỰC NGHIÊN CỨU 15
2.4 HIỂN THỊ RANH GIỚI LÊN GEE 15

CHƯƠNG 3 THỰC NGHIỆM PHÂN LOẠI CÓ GIÁM SÁT **16**

3.1 QUY TRÌNH PHÂN LOẠI 16
3.2 CÁC BƯỚC THỰC HIỆN 16
3.2.1 Bước 1 - Thu thập dữ liệu huấn luyện 16
3.2.2 Bước 2 – Khởi tạo trình phân loại 21
3.2.3 Bước 3 - Huấn luyện để tạo công cụ phân loại 24
3.2.4 Bước 4 – Phân loại hình ảnh 24
3.3 XỬ LÝ ẢNH SAU PHÂN LOẠI 25
3.3.1 Xác định và cập nhật các vùng 25
3.3.2 Chuyển đổi từ raster sang vector 25
3.3.3 Hiển thị kết quả vector 26
3.3.4 Run và kiểm tra kết quả 26
3.3.5 Xuất kết quả ra file theo định dạng GeoJSON 26
3.3.6 Tải file kết quả vào Google Drive của người dùng 27

TÀI LIỆU THAM KHẢO **28**

GIỚI THIỆU

Báo cáo nghiên cứu này cung cấp cho người dùng cuối (end-user) cách thức tham gia để trở thành thành Nhà phát triển làm việc với nền tảng đám mây Google Earth Engine. Sau khi trở thành Nhà phát triển, người dùng sẽ có các kỹ năng thực tế để có thể làm chủ nền tảng và thực hiện các dự án của cá nhân cũng như áp dụng trong giảng dạy về lĩnh vực xử lý dữ liệu không gian địa lý và hình ảnh vệ tinh quy mô lớn. Nghiên cứu này có thể làm tài liệu tham khảo giúp các giảng viên và trợ giảng có thể thiết kế các bài thực hành theo mục tiêu tác nghiệp.

Mục tiêu tác nghiệp của một người được hiểu là hành động của họ để hoàn thành công việc được giao theo vị trí việc làm. Ví dụ, đối với vị trí "người thực hành" hay "sinh viên làm bài thực hành" thì mục tiêu tác nghiệp của họ là thực hiện bài thực hành theo từng bước (bắt chước, làm theo) để đạt được kết quả đã đề ra trong bài thực hành đó, giúp cho sinh viên chiếm lĩnh kiến thức, hình thành kỹ năng nghề nghiệp, tự kiểm tra, điều chỉnh trong quá trình học; đối với vị trí "người hướng dẫn" hay "trợ giảng" thì mục tiêu tác nghiệp là hướng dẫn, tư vấn, hỗ trợ người học khi có tình huống phát sinh để họ hoàn thành bài thực hành theo yêu cầu.

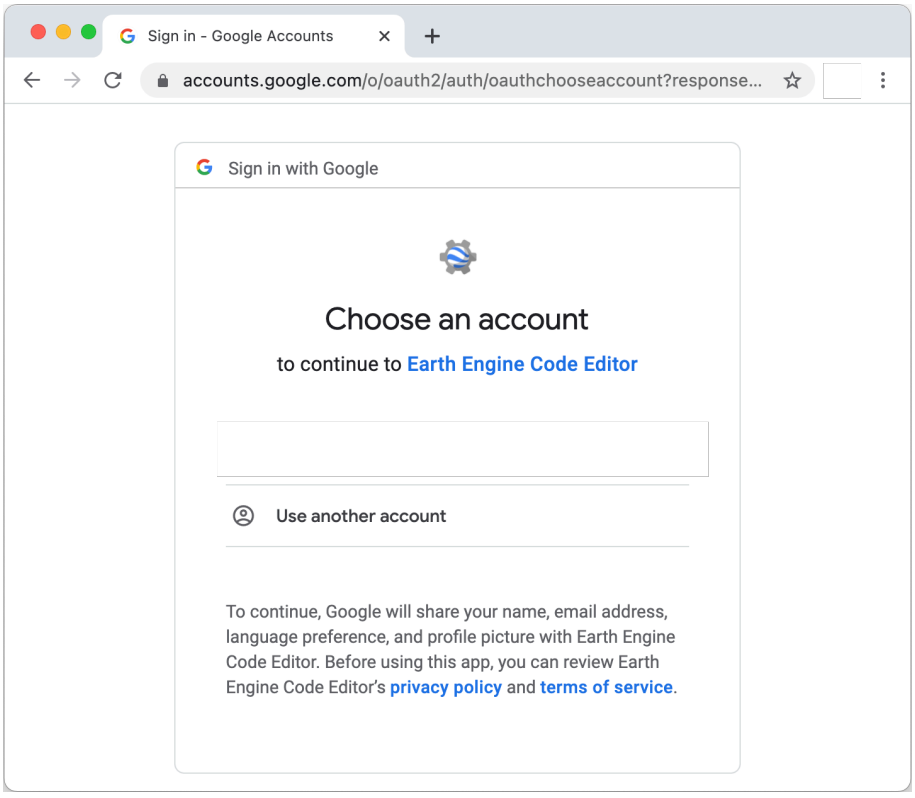
Do vậy, đối với vị trí việc làm là "người hướng dẫn" hay "trợ giảng" ngoài kỹ năng nghề nghiệp, họ cần phải am hiểu bài thực hành để giải quyết và xử lý các tình huống phát sinh trong khi hướng dẫn thực hành. Đối với vị trí người học, đây được cho là hướng tích cực hóa hoạt động học của sinh viên.

Chương 1
GIỚI THIỆU NỀN TẢNG ĐÁM MÂY GOOGLE EARTH ENGINE

Google Earth Engine (GEE, 2021) là một nền tảng dựa trên đám mây cho phép xử lý dữ liệu không gian, hình ảnh vệ tinh quy mô lớn. GEE cung cấp nền tảng tương tác giúp người dùng phát triển thuật toán xử lý dữ liệu không gian địa lý quy mô lớn để phát hiện các thay đổi, xu hướng bản đồ và định lượng sự khác biệt trên bề mặt Trái đất. Chương này cung cấp cho người dùng cách thức tham gia trở thành Nhà phát triển làm việc với nền tảng GEE.

1.1 Tạo tài khoản để làm việc với GEE

Truy cập vào liên kết *http://code.earthengine.google.com* bằng trình duyệt Google Chrome. *(Nếu một người dùng đã có tài khoản GEE thì có thể đăng nhập tài khoản để làm việc với GEE và bỏ qua bước này).* Một người dùng phổ thông muốn tham gia làm việc với GEE cần phải đăng ký bằng cách sử dụng tài khoản email sẵn có (nên dùng email của google @gmail.com) để liên kết với GEE.



Hình 1-1 Tạo tài khoản hoặc cấp quyền tài khoản sẵn có cho Google

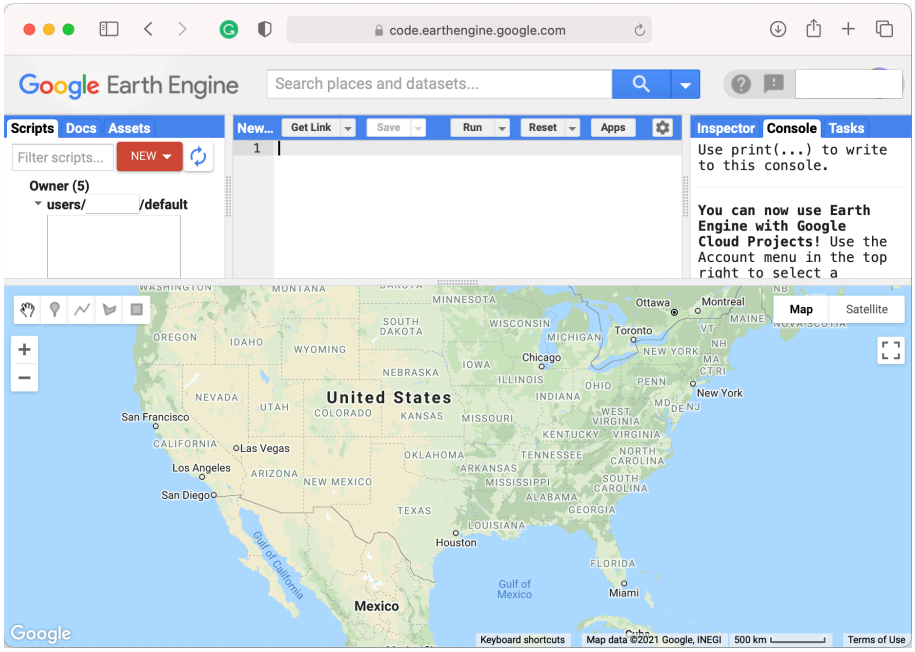
Cấp quyền truy cập tài khoản email cho GEE. Cài đặt mặc định sẵn có cũng có thể làm việc được với GEE.

1.2 Đăng nhập để làm việc với GEE

Nếu việc tạo tài khoản thành công hoặc đã có tài khoản sẵn có thì tiến hành đăng nhập bằng cách truy cập vào liên kết *http://code.earthengine.google.com*

Nếu thành công, người dùng sẽ có một kho lưu trữ mã nguồn mới có tên **users/<tên>/default** trong tab Scripts trong phần Owner.

Kết quả hiển thị: Một cửa sổ trình soạn thảo mã trông giống như bên dưới



Hình 1-2 Giao diện làm việc của Nhà phát triển

1.3 Giới thiệu trình soạn thảo mã của GEE

- Tab **Scripts** là khu vực chứa các thư mục và các tệp mã nguồn, tại đây người dùng có thể tạo file để lưu mã nguồn cho những lần mở sau.
- Tab **Docs** là khu vực chứa tài liệu, tại đây người dùng có thể tra cứu các hàm và các tham số của hàm để biết cách truyền đối số cho phù hợp.

- Tab **Assets** là khu vực chứa tài nguyên của người dùng, tại đây người dùng có thể tạo tài nguyên và sử dụng tài nguyên cho việc viết mã.
- Tab **Inspector** là khu vực kiểm soát thông tin, tại đây người dùng có thể sử dụng chuột nhấp vào bản đồ để xem thông tin.
- Tab **Console** là khu vực hiển thị kết quả kiểm tra trong quá trình thực thi mã.
- Tab **Tasks** là khu vực hiển thị các tác vụ mà người dùng tạo ra như tải lên các dữ liệu lên hoặc tải xuống các kết quả.
- Một cửa sổ ở giữa dùng để viết mã hoặc hiển thị mã, đây là khu vực chính của người dùng để viết mã và làm việc với GEE.
- Nút **Get Link** có chức năng tạo ra một liên kết để chia sẻ cho cộng sự phiên bản hiện tại của mã.
- Nút **Save** có chức năng lưu lại mã hiện tại vào file và quản lý chúng trong tab Scripts cho những lần mở tiếp theo.
- Nút **Run** có chức năng thực thi mã hiện tại ở cửa sổ soạn thảo mã, kết quả có thể hiển thị ở tab Console hoặc khu vực bản đồ và ảnh ở phía dưới.
- Nút **Reset** có chức năng đặt lại trạng thái để tiến hành thực thi.

Nút **Apps** có chức năng tạo các ứng dụng thực nghiệm với Google Cloud Platform (GCP), người dùng cần phải thanh toán tiền để được sử dụng GCP.

1.4 Học máy trong GEE

Học máy (ML) trong Earth Engine được hỗ trợ để huấn luyện dữ liệu tạo ra mô hình học máy nhằm suy luận hình ảnh vệ tinh.

Có thể tìm thấy các ví dụ này ở trang https://developers.google.com/earth-engine/guides/classification (GEE, Supervised Classification, 2021) và trang https://developers.google.com/earth-engine/guides/clustering (GEE, Unsupervised Classification, 2021) tương ứng với Phân loại có giám sát và Phân loại không giám sát.

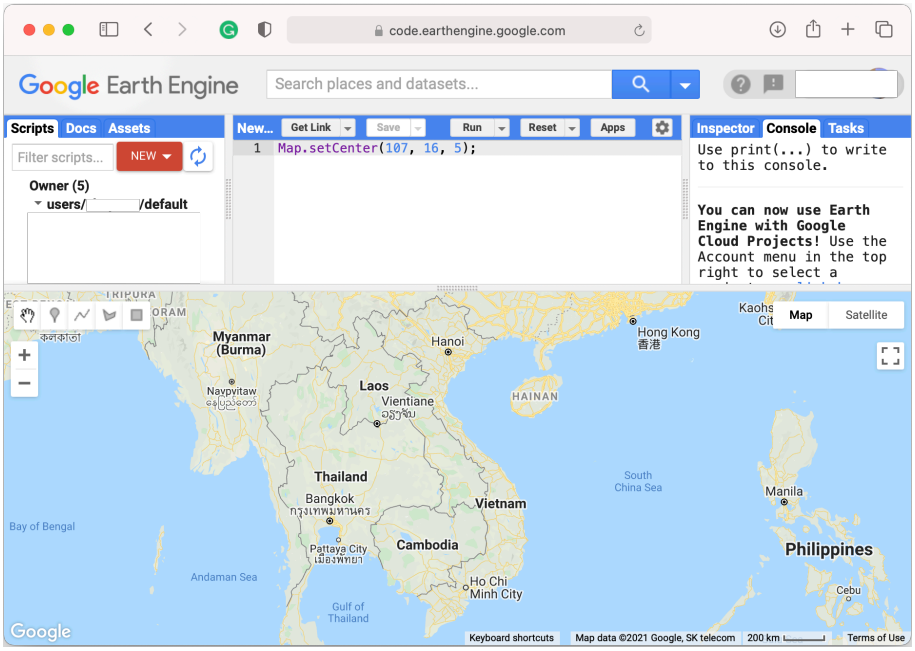
Chương 2
MỘT SỐ CHỨC NĂNG CƠ BẢN CỦA GOOGLE EARTH ENGINE

Với GEE, các nhà phát triển có thể xem hình ảnh vệ tinh, bản đồ, địa hình, tòa nhà 3D, v.v. Mặt khác, GEE là một công cụ để phân tích thông tin không gian địa lý. Các nhà nghiên cứu có thể phân tích độ che phủ của rừng và nước, sự thay đổi sử dụng đất hoặc đánh giá hiệu quả của các lĩnh vực nông nghiệp, và nhiều phân tích khả thi khác.

2.1 Căn giữa khu vực đã biết tọa độ và mức thu phóng

- Lệnh: Map.setCenter(lon, lat, zoom)
- Các đối số:
    - lon (số, bắt buộc): kinh độ điểm giữa của khu vực, đơn vị là độ.
    - lat (số, bắt buộc): vĩ độ điểm giữa của khu vực, đơn vị là độ.
    - zoom (số, tùy ý): mức thu phóng, từ 1 đến 24.
- Kết quả trả về: Giao diện bản đồ

Ví dụ: Map.setCenter(107, 16, 5);



Hình 2-1 Kết quả căn giữa và mức thu phóng

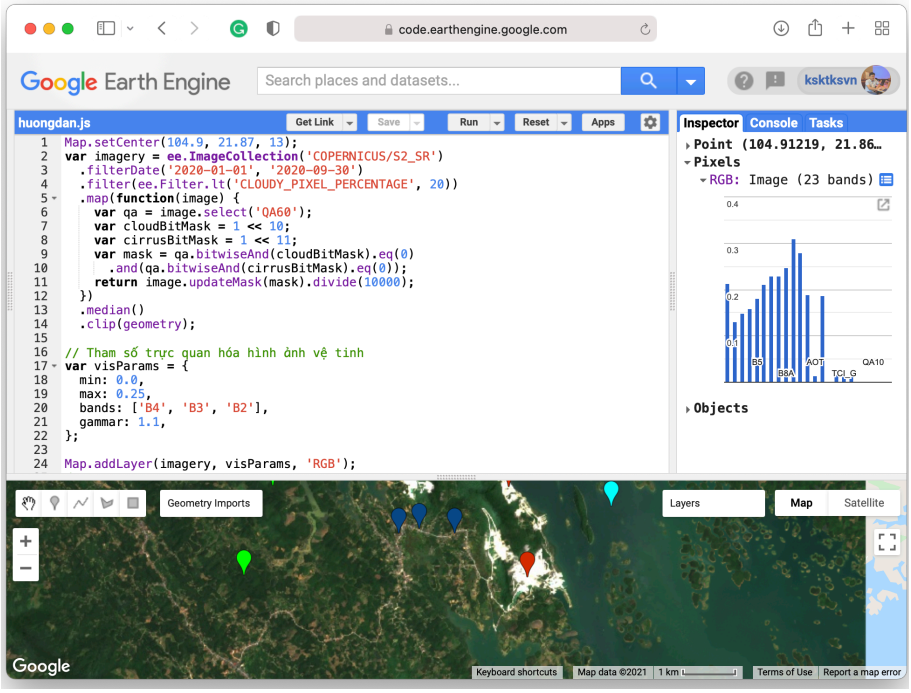
2.2 Thêm lớp dữ liệu ảnh Sentinel-2

- Các lệnh: Công việc này đòi hỏi thực hiện nhiều lệnh (tập lệnh)
- Kết quả trả về: Giao diện bản đồ có chứa layer ảnh Sentinel-2

Ví dụ:

```
Map.setCenter(104.9, 21.87, 13);
var imagery = ee.ImageCollection('COPERNICUS/S2_SR')
  .filterDate('2020-01-01', '2020-09-30')
  .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 20))
  .map(function(image) {
    var qa = image.select('QA60');
    var cloudBitMask = 1 << 10;
    var cirrusBitMask = 1 << 11;
    var mask = qa.bitwiseAnd(cloudBitMask).eq(0)
      .and(qa.bitwiseAnd(cirrusBitMask).eq(0));
    return image.updateMask(mask).divide(10000);
  });
var visParams = {
  min: 0.0,
  max: 0.3,
  bands: ['B4', 'B3', 'B2'],
  gammar: 1.1,
};
Map.addLayer(imagery.median(), visParams, 'RGB');
```

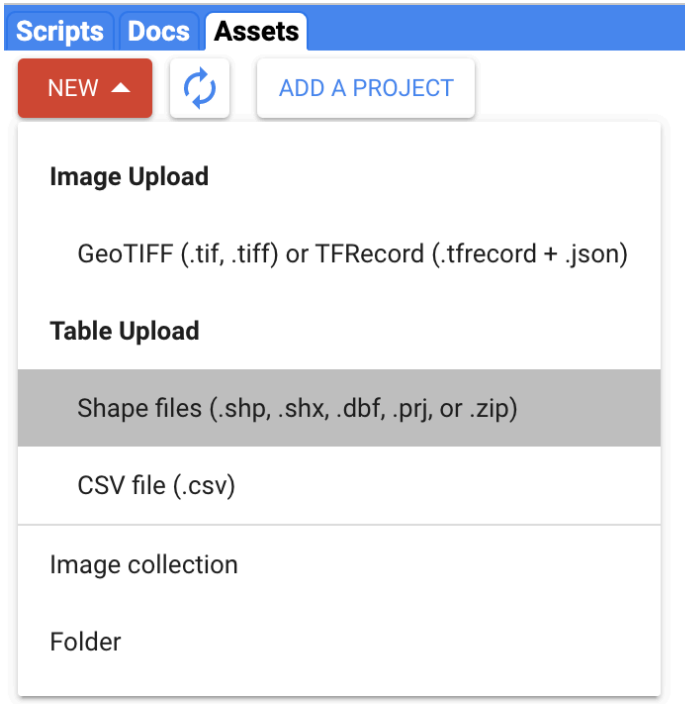
Xác nhận hình ảnh thực hiện:



Hình 2-2 Kết quả thêm lớp dữ liệu ảnh Sentinel-2

2.3 Nhập ranh giới khu vực nghiên cứu

Mở Tab **Assets**, nhấn vào nút **NEW**, sau đó chọn **Shape files** ở trong phần **Table Upload.**



Hộp thoại như hình ở dưới xuất hiện. Nhấn vào **SELECT** và chọn file Shapefile hoặc file ZIP chứa Shapefile trên máy tính. Khi chọn tệp .shp, cần chọn các tệp .dbf, .shx và .prj. Lưu ý rằng các tệp .shp, .dbf và .shx là bắt buộc. Các tệp khác là tùy chọn. Nếu tệp .prj không được cung cấp, Earth Engine sẻ mặc định sử dụng WGS84.

2.4 Hiển thị ranh giới lên GEE

Chương 3
THỰC NGHIỆM PHÂN LOẠI CÓ GIÁM SÁT

Gói Classifier xử lý phân loại có giám sát bằng các thuật toán ML truyền thống chạy trong Earth Engine. Các bộ phân loại này bao gồm CART, RandomForest, NaiveBayes và SVM.

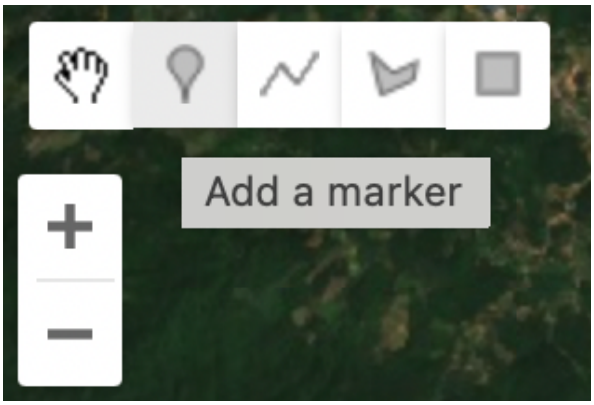
3.1 Quy trình phân loại

1. Thu thập dữ liệu huấn luyện. Tập hợp các đối tượng có thuộc tính và lưu chúng vào một lớp với nhãn đã biết và thuộc tính lưu trữ các giá trị số cho các yếu tố dự đoán.
2. Khởi tạo trình phân loại. Đặt các thông số của nó nếu cần thiết.
3. Huấn luyện để tạo công cụ phân loại (tạo ra mô hình ML) bằng cách sử dụng dữ liệu huấn luyện ở bước 1.
4. Phân loại hình ảnh hoặc bộ sưu tập các đối tượng địa lý.

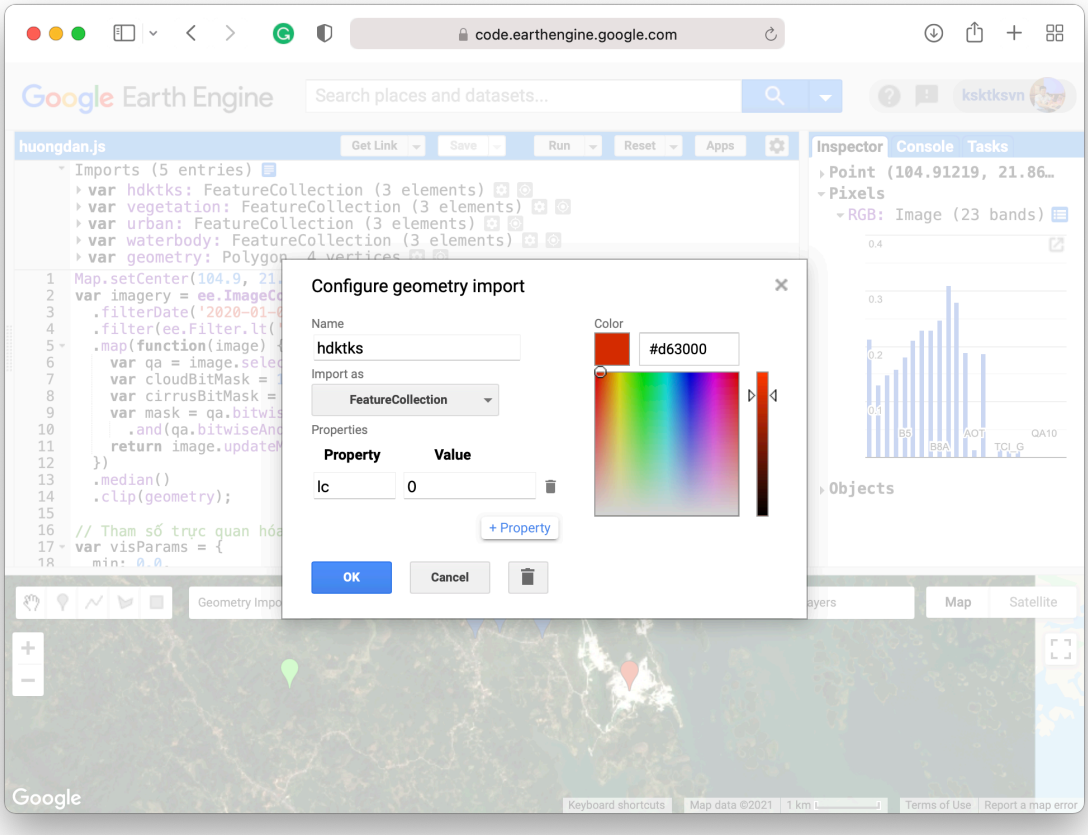
3.2 Các bước thực hiện

3.2.1 Bước 1 - Thu thập dữ liệu huấn luyện

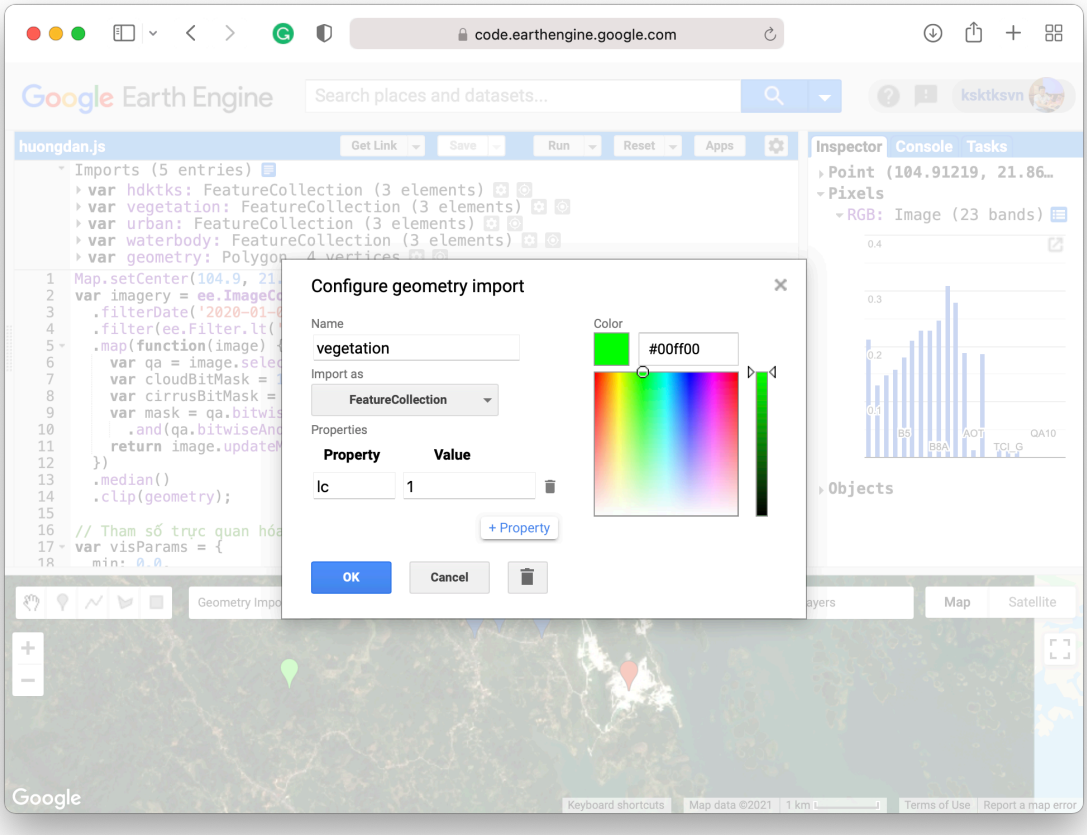
Dữ liệu huấn luyện là một FeatureCollection với một thuộc tính lưu trữ nhãn lớp và các thuộc tính lưu trữ các biến dự đoán. Sử dụng công cụ lấy mẫu có sẵn của GEE để tiến hành thu thập dữ liệu huấn luyện.



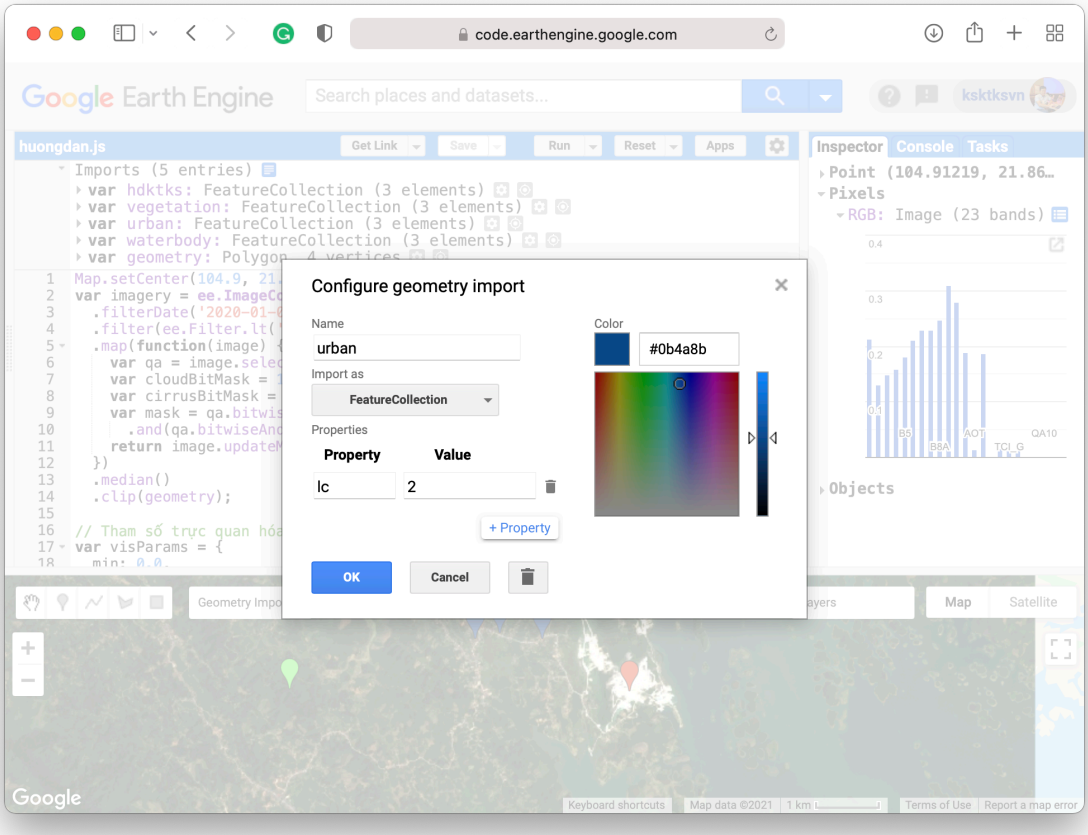
- Chuẩn bị lớp phủ Hoạt động khai thác khoáng sản **“hdktks”**



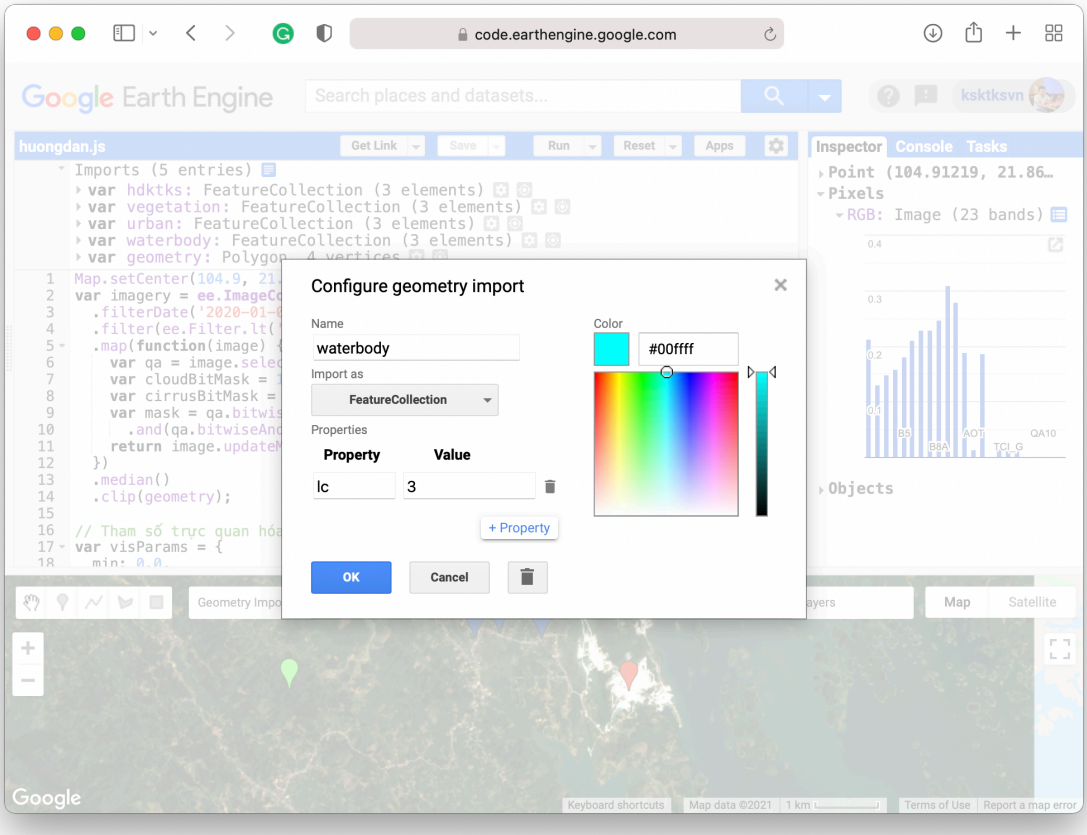
- Chuẩn bị lớp phủ Thực vật **“vegetaion”**



- Chuẩn bị lớp phủ Dân cư, đô thị **“urban”**

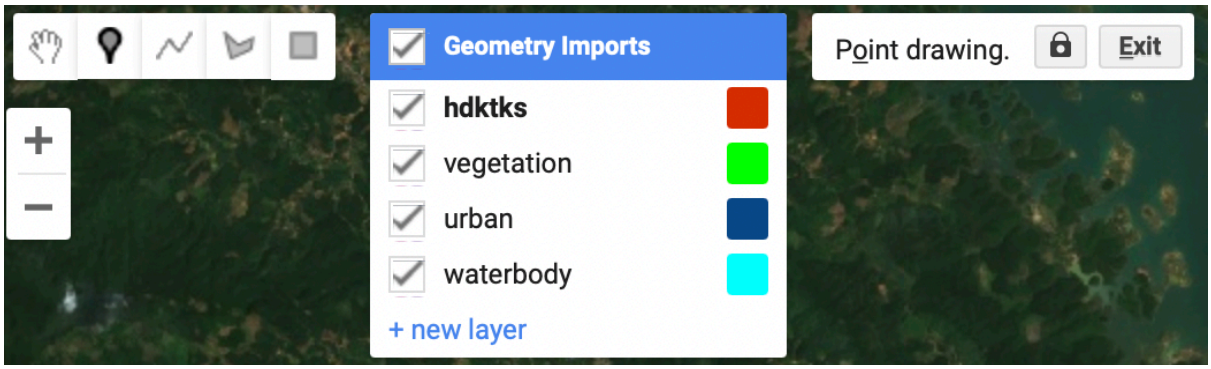


- Chuẩn bị lớp phủ Nước mặt **“waterbody”**



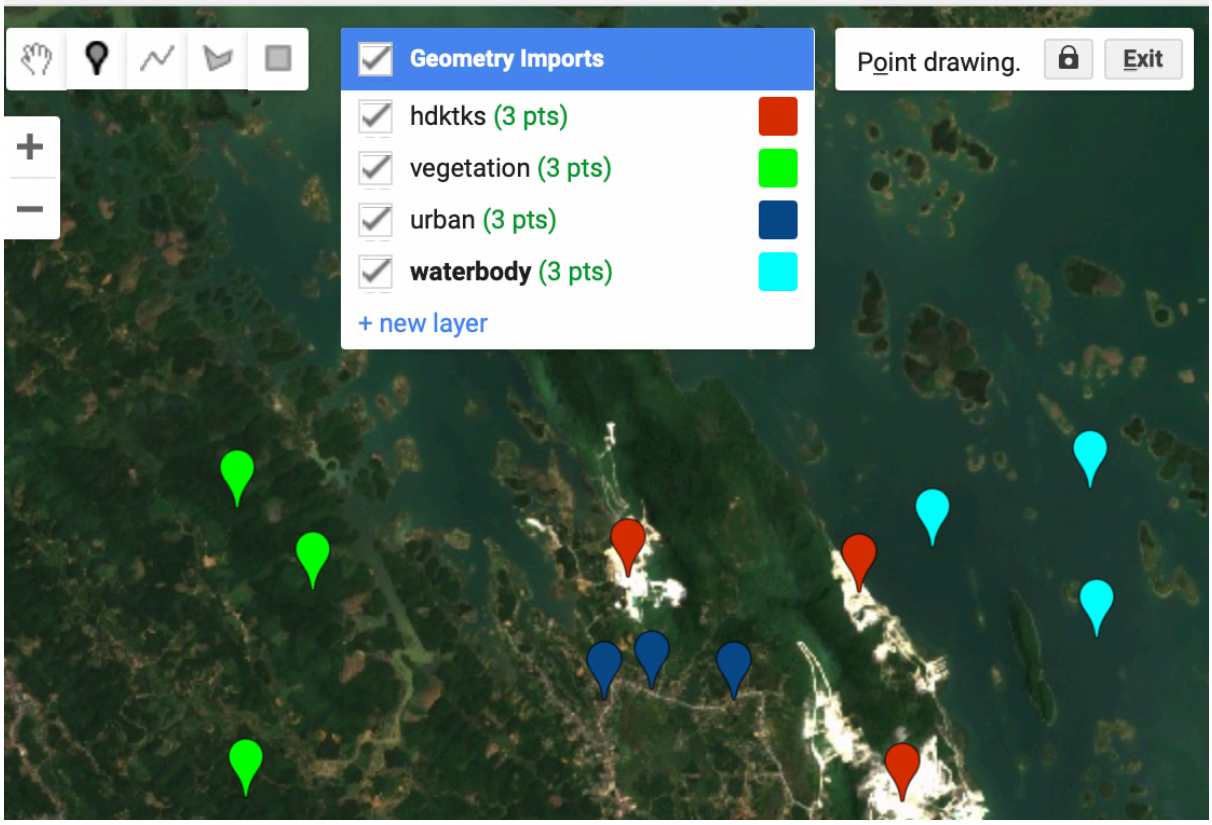
Tại thanh công cụ của GEE:

- Lựa chọn lớp phủ bằng cách click vào **Geometry Imports** và chọn, lớp mặc định sẽ được tô đậm
- Chọn kiểu dữ liệu mẫu bằng cách chọn **Add a marker**
- Di chuyển đến giữa vùng mẫu để đặt điểm
- Thực hiện các bước tương tự cho tất cả các lớp, dữ liệu thu thập được sẽ tự hiển thị ở phần Imports của file mã nguồn.



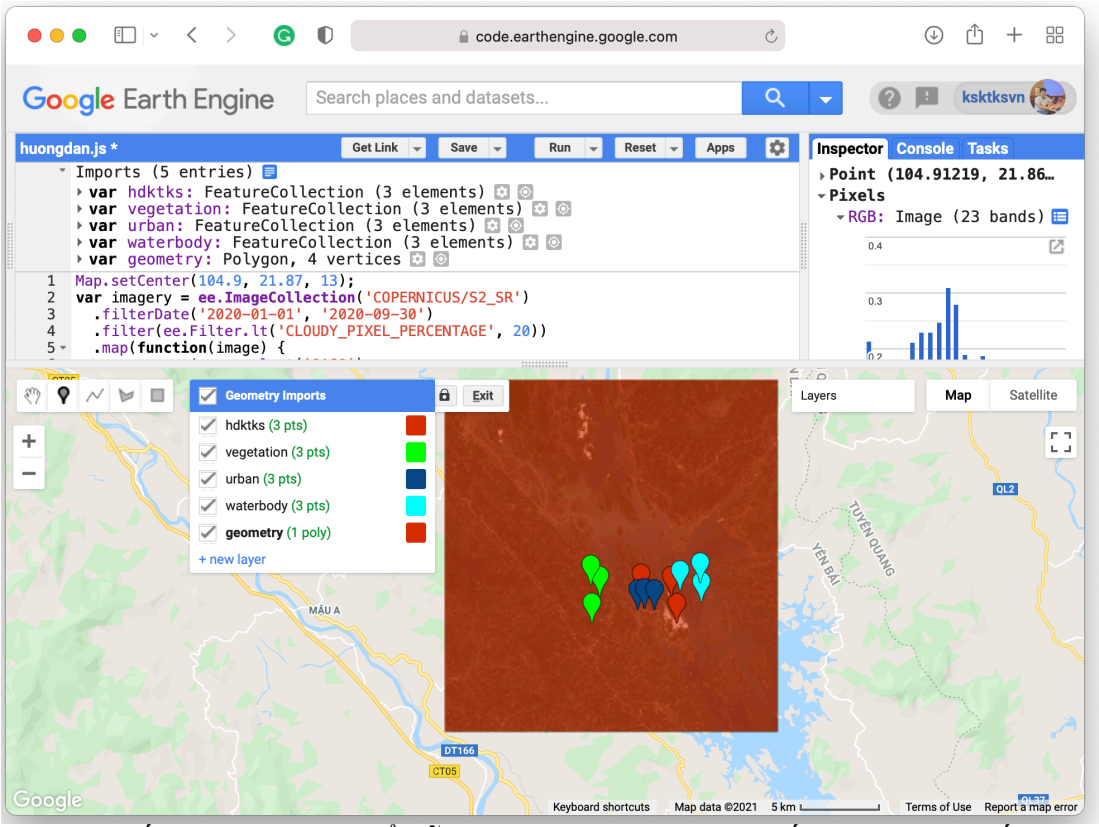
Với số lượng mẫu đủ lớn và chính xác thì tiến hành khởi tạo trình phân loại ở bước 2

```
New Script *
Imports (4 entries)
  var hdktks: FeatureCollection (3 elements)
  var vegetation: FeatureCollection (3 elements)
  var urban: FeatureCollection (3 elements)
  var waterbody: FeatureCollection (3 elements)
1  Map.setCenter(104.9, 21.87, 13);
2  var dataset = ee.ImageCollection('COPERNICUS/S2_SR')
3    .filterDate('2020-01-01', '2020-09-30')
4    .filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 20))
5    .map(function(image) {
6      var qa = image.select('QA60');
7      var cloudBitMask = 1 << 10;
8      var cirrusBitMask = 1 << 11;
```

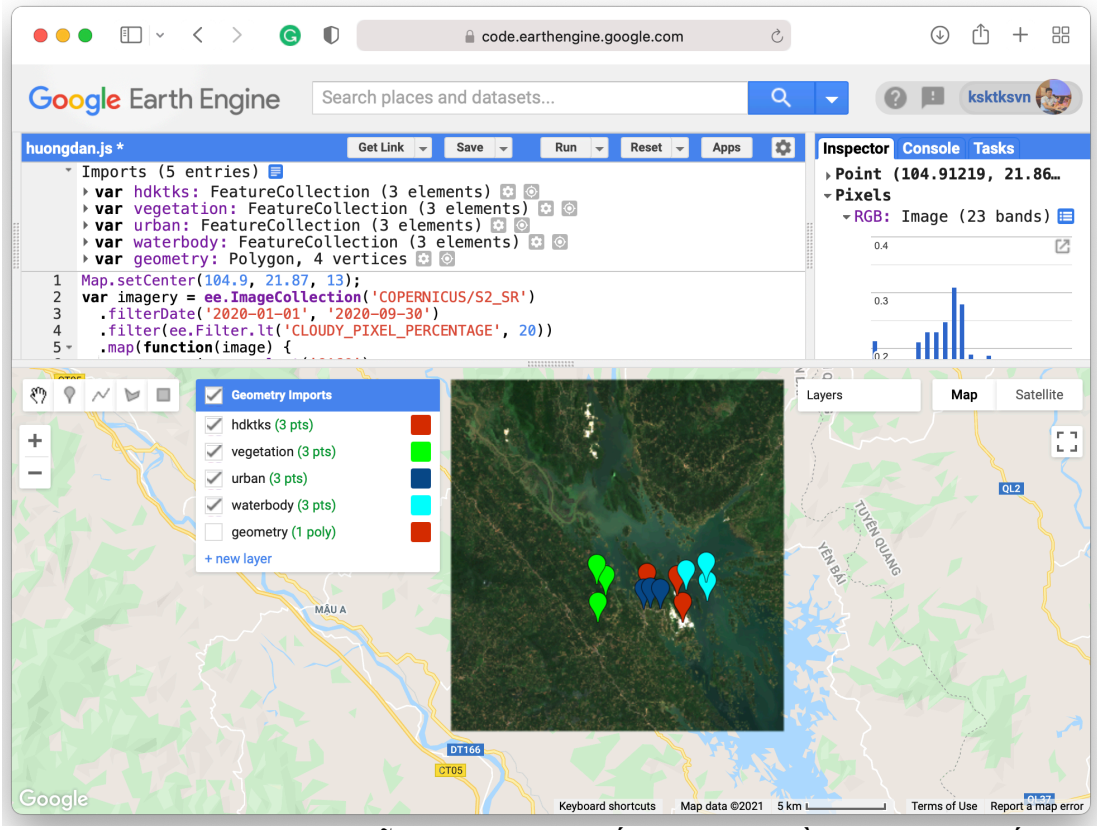


3.2.2 Bước 2 – Khởi tạo trình phân loại

Tạo một khu vực nghiên cứu có tên "**geometry**" bằng cách **add new layer**, sau đó dùng công cụ vẽ hình chữ nhật để tạo. Sau đó tắt lớp geometry để dễ quan sát.



Chú ý: Nên tắt lớp geometry để dễ quan sát. Run và được kết quả mới đã cắt vùng nghiên cứu:



Tạo một tập dữ liệu mẫu và lưu vào biến **samples** bằng cách gộp tất cả các mẫu được lấy trước đó.

```
// Tạo tập dữ liệu mẫu
var samples = hdktks.merge(vegetation).merge(urban).merge(waterbody);
```

Tạo nhãn lưu giữ thuộc tính lớp phủ dưới dạng các số nguyên liên tiếp bắt đầu từ số không.

```
// Tạo nhãn lưu giữ thuộc tính lớp phủ dưới dạng các số nguyên liên tiếp
bắt đầu từ số không.
var label = 'lc';
```

Sử dụng các kênh của ảnh để dự đoán. Để có được kết quả tốt, phù hợp với từng bài toán thì cần phải tổ hợp thêm kênh.

```
// Sử dụng các kênh của ảnh để dự đoán.
// Để có được kết quả tốt, phù hợp với từng bài toán thì cần phải tổ hợp
thêm kênh.
var bands = ['B2', 'B3', 'B4', 'B5', 'B6', 'B7', 'B10', 'B11'];
```

Chồng dữ liệu mẫu lên hình ảnh để chuẩn bị huấn luyện. Đây là quá trình chuyển đổi từng pixel của hình ảnh (ở một tỷ lệ nhất định) thành đối tượng địa lý, kết quả trả về là một FeatureCollection. Mỗi đối tượng đầu ra sẽ có một thuộc tính trên mỗi kênh của hình ảnh đầu vào, cũng như bất kỳ thuộc tính cụ thể nào được sao chép từ đối tượng đầu vào.

Lưu ý rằng vị trí hình học của đối tượng sẽ được gắn vào tâm pixel.

```
// Chồng dữ liệu mẫu lên hình ảnh để chuẩn bị huấn luyện.
var trainSet = imagery.select(bands).sampleRegions({
    collection: samples,
    properties: [label],
    scale: 30
});
```

3.2.3 Bước 3 - Huấn luyện để tạo công cụ phân loại

Tạo một công cụ phân loại theo Random Forest.

```
// Huấn luyện để tạo công cụ phân loại theo các
var mlModel = ee.Classifier.smileRandomForest(10).train(trainSet, label,
bands);
```

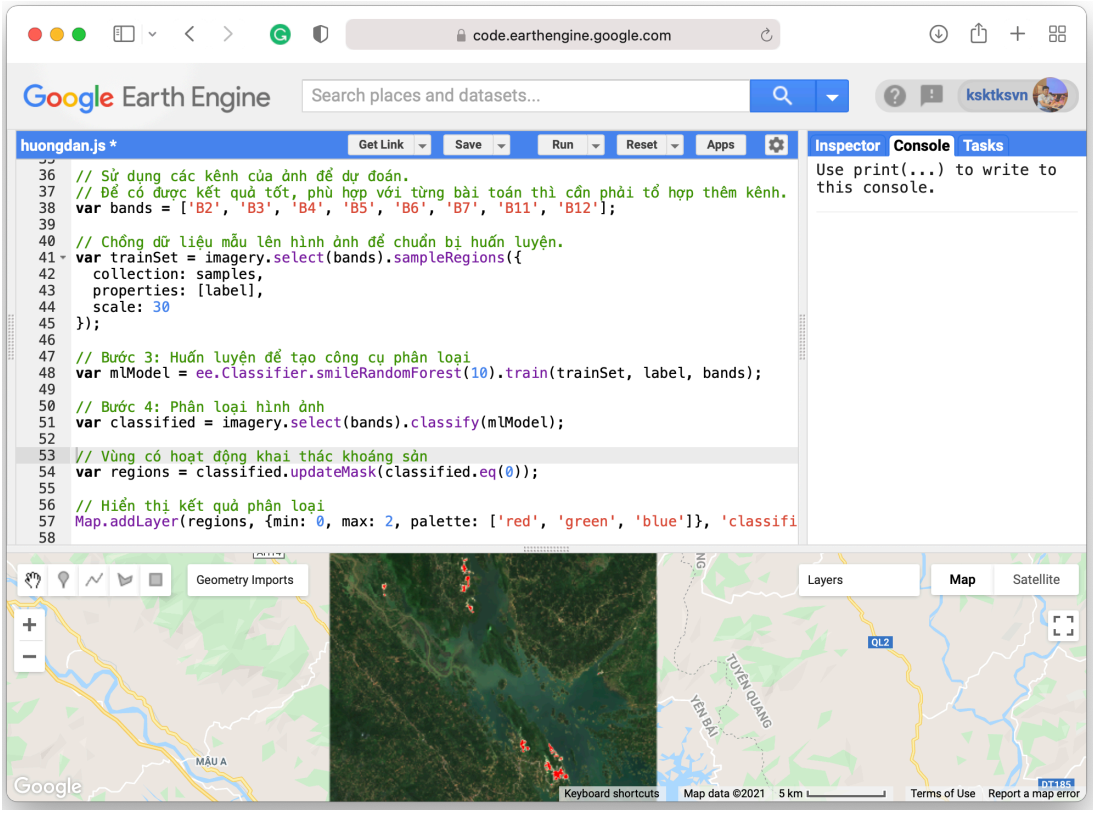
3.2.4 Bước 4 – Phân loại hình ảnh

```
// Phân loại hình ảnh với các kênh như đã được sử dụng để huấn luyện.
var classified = imagery.select(bands).classify(mlModel);

// Vùng có hoạt động khai thác khoáng sản
var regions = classified.updateMask(classified.eq(0));

// Hiển thị kết quả phân loại
Map.addLayer(regions, {min: 0, max: 2, palette: ['red', 'green', 'blue']},
'classification');
```

Run và cho kết quả:



3.3 Xử lý ảnh sau phân loại

Sau khi có hình ảnh kết quả phân loại, cần phải tiến hành xử lý để có được kết quả cuối cùng và xuất dữ liệu ra file vector.

3.3.1 Xác định và cập nhật các vùng

```
var minArea = 2500;
var maxSize = 20;
var pixelCount = classified.connectedPixelCount(maxSize);
var minPixelCount = ee.Image(minArea).divide(ee.Image.pixelArea());
var classifiedFin = classified.updateMask(pixelCount.gte(minPixelCount));
```

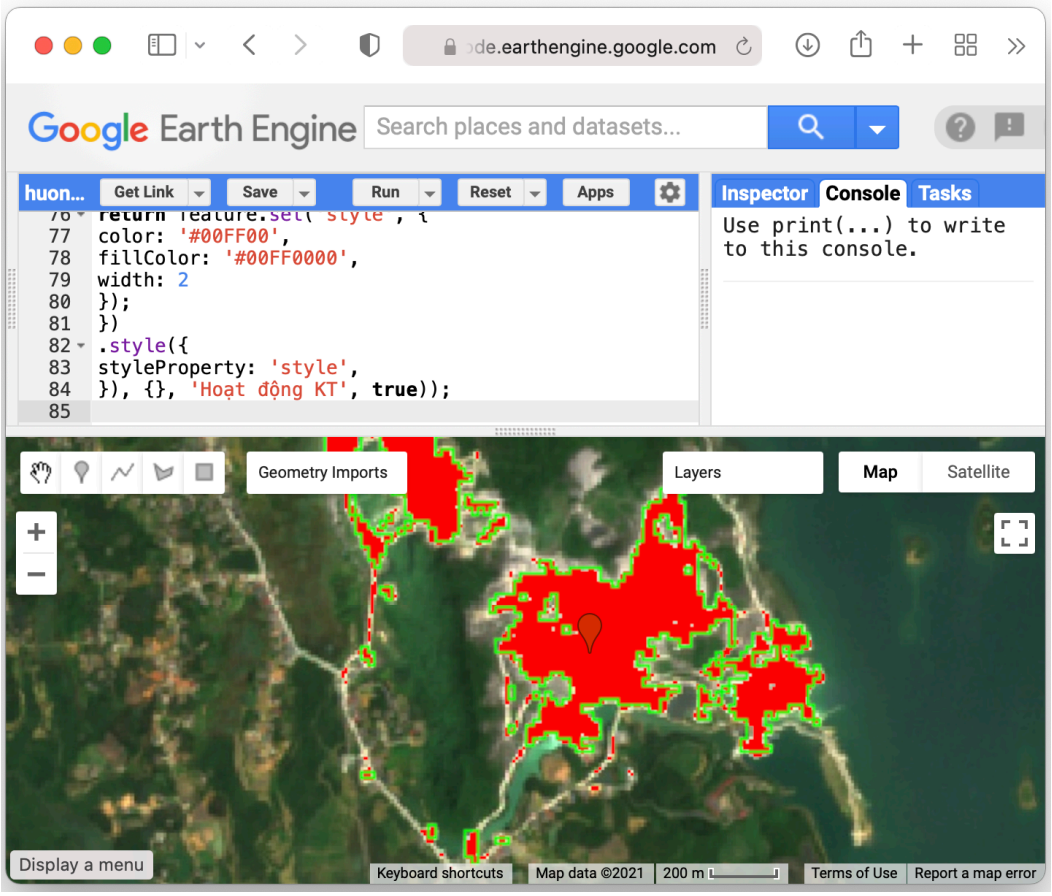
3.3.2 Chuyển đổi từ raster sang vector

```
var resultVector = regions.reduceToVectors({
reducer: ee.Reducer.countEvery(),
geometry: geometry,
scale: 20,
geometryInNativeProjection: true,
maxPixels: 1e13,
tileScale: 4
});
```

3.3.3 *Hiển thị kết quả vector*

```
Map.layers().set(3, ui.Map.Layer(resultVector.map(function (feature) {
return feature.set('style', {
color: '#00FF00',
fillColor: '#00FF0000',
width: 2
});
})
.style({
styleProperty: 'style',
}), {}, 'Hoạt động KT', true));
```

3.3.4 *Run và kiểm tra kết quả*

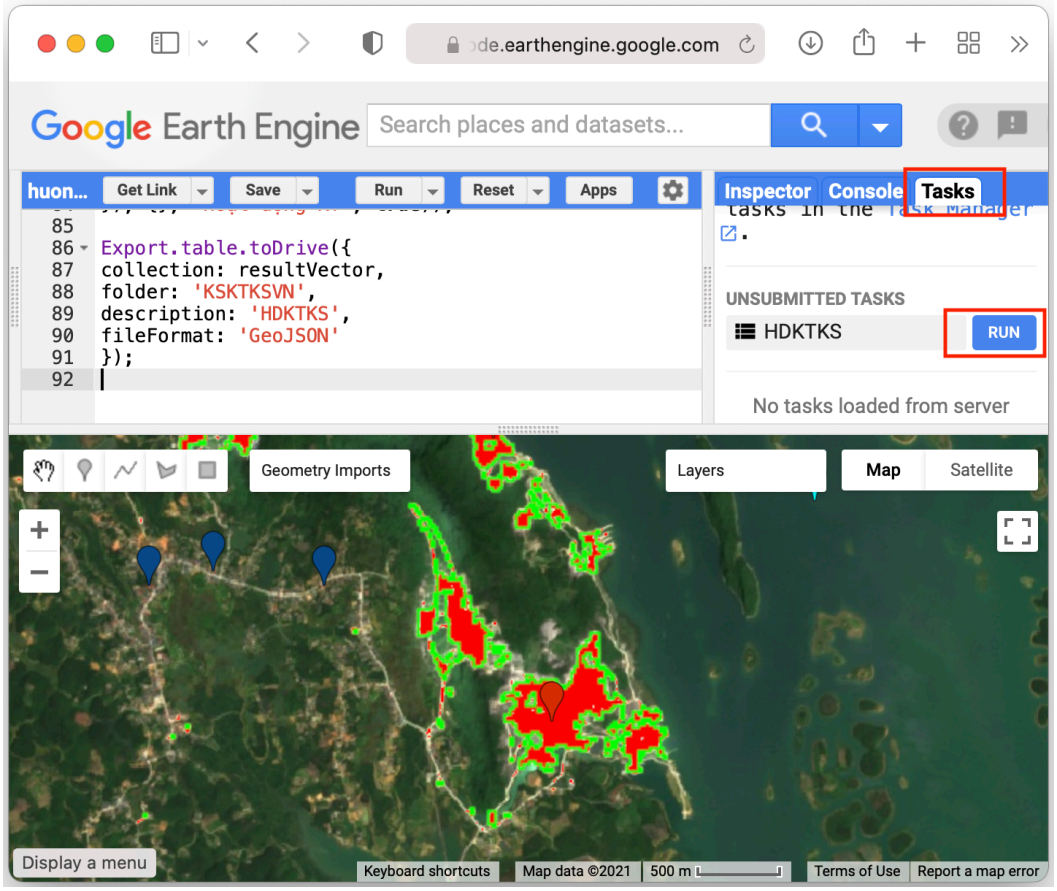


3.3.5 *Xuất kết quả ra file theo định dạng GeoJSON*

```
Export.table.toDrive({
collection: resultVector,
folder: 'KSKTKSVN',
description: 'HDKTKS',
fileFormat: 'GeoJSON'
});
```

3.3.6 Tải file kết quả vào Google Drive của người dùng

Tiến hành Run để xuất kết quả, sau đó chuyển qua tab **Tasks** để tải file bằng cách nhấn vào nút **RUN**. Quá trình tải file kết thúc thì người dùng có thể vào thư mục trên Drive của mình để tải về máy tính.



TÀI LIỆU THAM KHẢO

GEE. (2021). *Google Earth Engine*. Retrieved 12, 2021 from https://developers.google.com/earth-engine

GEE. (2021). *Supervised Classification*. Retrieved 12, 2021 from Google Earth Engine: https://developers.google.com/earth-engine/guides/classification

GEE. (2021). *Unsupervised Classification*. Retrieved 12, 2021 from Google Earth Engine: https://developers.google.com/earth-engine/guides/clustering