

RAISG

RED AMAZÓNICA DE INFORMACIÓN  
SOCIOAMBIENTAL GEORREFERENCIADA



MAPBIOMAS  
[AMAZONIA]

Apéndice - Minería

Tema transversal en Colección 6

Versión 1

Equipo responsable  
Emanuel Valero  
Rodrigo Lazo  
José Sánchez

Septiembre 2024

# Índice

<b>1. Introducción</b>	<b>3</b>
<b>2. Construcción y definición de mosaicos Landsat</b>	<b>3</b>
<b>3. Variables de clasificación</b>	<b>5</b>
<b>4. Clasificación</b>	<b>5</b>
4.1. Búsqueda y actualización de datos	5
4.2. Preparación de los datos	6
4.3. Muestras de entrenamiento (minería/no minería)	6
4.4. Clasificación con algoritmo Random Forests	7
<b>5. Post-clasificación</b>	<b>8</b>
5.1 Filtro espacial	8
5.2 Filtro de frecuencia	8
5.3 Integración	9
<b>6. Resumen metodológico</b>	<b>11</b>
<b>7. Referencias</b>	<b>11</b>

# Lista de figuras

<i>Figura 1: Área de estudio</i>	4
<i>Figura 2: Tabla de coordenadas y mapa de puntos con la localización de las zonas mineras</i>	6
<i>Figura 3: Polígono de zonas mineras</i>	6
<i>Figura 4: Polígono de Mina y No-Mina</i>	7
<i>Figura 5: Der. Mina de bauxita. Izq. Clasificación de zona minera</i>	8
<i>Figura 6: Esquema de la funcionalidad del filtro espacial</i>	8
<i>Figura 7: Efecto del filtro de frecuencia</i>	9
<i>Figura 8: Clasificación de zona minera integrada</i>	9
<i>Figura 9: Imágenes de referencia (resultado de la integración del tema transversal de minería en el mapa general, imagen de alta resolución de Google Earth y paisaje)</i>	10
<i>Figura 10: Pasos para la clasificación de minería</i>	11

# Lista de Tablas

<i>Tabla 1: Variables para la clasificación de zonas mineras de la Colección 6</i>	5
--	---

## 1. Introducción

La identificación de áreas mineras se considera un tema transversal dentro de la metodología de MapBiomias, ya que puede generar confusiones al clasificar diferentes tipos de cobertura del suelo, como herbazales, pastos o zonas agrícolas. Estas áreas suelen presentar características visuales similares, dificultando una diferenciación precisa. Por ello se ha desarrollado una metodología particular para poder clasificar el uso minero en las imágenes de satélite de Landsat. En la colección 6, se incluyeron los siguientes países: Bolivia, Colombia, Ecuador, Guyana, Guayana Francesa, Perú, Surinam y Venezuela.

Este documento detalla la metodología empleada en la colección 6 para clasificar zonas mineras, las cuales abarcan todas las áreas de extracción de materiales pétreos o minerales con exposición clara del suelo. No se hace distinción entre los tipos de actividades mineras, ya sean industriales, artesanales, legales o ilegales.

Cabe destacar que la metodología original se basó en la experiencia del equipo de MapBiomias Brasil, especialmente en el trabajo de César Diniz (Solutions in Geoinformation, Solved, Brasil). Diniz presentó al equipo de MapBiomias las técnicas utilizadas para identificar áreas de minería industrial en Brasil. Estas metodologías fueron adaptadas y mejoradas en la plataforma Google Earth Engine, con ajustes menores para optimizar tanto el rendimiento como la calidad de los resultados.

## 2. Construcción y definición de mosaicos Landsat

Para el procesamiento de datos, la metodología de MapBiomias divide el área del proyecto en una cuadrícula de celdas regulares, basada en la grilla de las Cartas Internacionales del Mundo a la Millonésima, con una escala de 1:250.000. Cada celda rectangular cubre un área equivalente a 1°30' de longitud y 1° de latitud. En la Figura 1 se presenta el área de estudio de la RAISG, que abarca ocho países: Guayana Francesa, Surinam, Guyana, Venezuela, Colombia, Ecuador, Perú y Bolivia

En cada una de las celdas de la cuadrícula, se generan mosaicos anuales a partir de imágenes Landsat. Estos mosaicos tienen como objetivo producir imágenes periódicas libres de nubes, lo que facilita su clasificación y la identificación de cambios en el uso del suelo. Para obtener más detalles, se recomienda consultar la sección 3.1 del documento *ATBD General MapBiomias Amazonía 6.0*. En esta colección, se empleó la Landsat Collection 2, que representa la versión más reciente del reprocesamiento de los datos de Landsat, incorporando varias mejoras en la calidad de las imágenes.

La construcción de los mosaicos comienza con la selección de imágenes Landsat. Para ello, se elabora una tabla de parametrización que contiene los siguientes elementos, almacenados como metadatos en cada conjunto de datos generado:

- ID: Identificador o código único de la unidad carta/región
- Año: Año de la serie (comprendiendo el período 1985 a 2023) al que corresponde el mosaico.
- Carta: Código identificador de la carta.

- Fecha inicial/Fecha final: Periodo del año (fecha de inicio y final) para la selección de imágenes del catálogo de datos de imágenes Landsat de Google Earth Engine.
- Sensor: El satélite y su respectivo sensor: Landsat 5 TM, Landsat 7 ETM+, Landsat 8 OLI, Landsat 9 o una combinación de Landsat 5 y Landsat 7.
- Porcentaje de nubes: Porcentaje máximo de cobertura de nubes aceptado de cada imagen Landsat que será usado para construir el mosaico de imágenes. Este dato proviene de la metadata de las imágenes Landsat.
- Blacklist: Imágenes que por su calidad son excluidas de la construcción del mosaico.

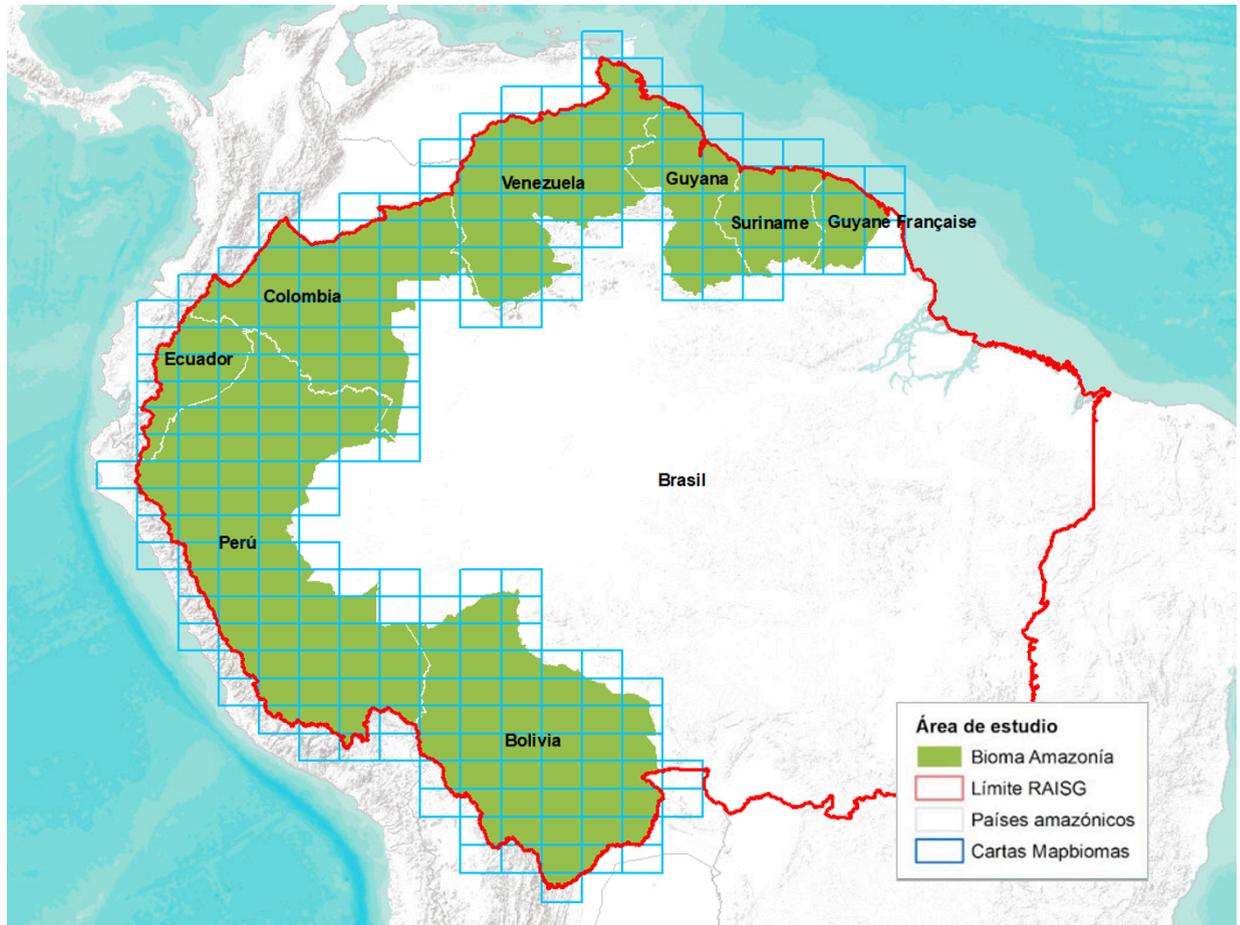


Figura 1: Área de estudio

Con esta información, se generan los mosaicos anuales de imágenes utilizando la herramienta Google Earth Engine. Todos los mosaicos se almacenan como una colección de imágenes en Google Earth Engine, disponibles en la ruta: [projects/mapbiomas-raisg/MOSAICOS/mosaics-2](https://earthengine.google.com/projects/mapbiomas-raisg/MOSAICOS/mosaics-2). Inicialmente, los mosaicos contienen los valores medianos de cada una de las bandas de los satélites Landsat. A partir de estos datos, se agregaron bandas adicionales que incluyen índices espectrales, fracciones espectrales, variables de textura y parámetros topográfico-geomorfológicos derivados del modelo digital de elevación.

A cada una de las bandas de los mosaicos se les aplica un análisis estadístico, generando información detallada sobre la mediana, la mediana de la estación seca (percentil 25), la mediana de la estación lluviosa (percentil 75), la amplitud (rango de valores), la desviación

estándar y el valor mínimo. Todos estos datos constituyen las variables que alimentan el clasificador Random Forest (bosque aleatorio), utilizado para el análisis final.

### 3. Variables de clasificación

La selección de variables derivadas de los mosaicos es el primer paso clave para la identificación de las zonas mineras en la Pan-Amazonía. A través de diversas pruebas, se logró identificar un conjunto específico de variables que optimizan la detección de actividades mineras. En total, se utilizaron 18 variables, las cuales se detallan en la Tabla 1.

Tabla 1 - Variables para la clasificación de zonas mineras de la Colección 6, MapBiomias Amazonia

Banda	Expresión / Descripción	Reductor
blue	Valor de la banda azul	Mediana
green	Valor de la banda verde	Mediana
red	Valor de la banda roja	Mediana
nir	Valor de la banda infrarrojo cercano (NIR)	Mediana
swir1	Valor de la banda SWIR1	Mediana
swir2	Valor de la banda SWIR2	Mediana
ndfi		Mediana
ndfib		Mediana
npv		Mediana
soil		Mediana, Amplitud
evi2	$2.5 * ((\text{NIR} - \text{Red}) / (\text{NIR} + 2.4 * \text{Red} + 1))$	Mediana
ndvi	$(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$	Mediana
ndwi_gao	$(\text{Green} - \text{SWIR1}) / (\text{Green} + \text{SWIR1})$	Mediana
ndwi_mcfeters	$(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$	Mediana
mndwi	$(\text{Green} - \text{NIR}) / (\text{Green} + \text{NIR})$	Mediana
gcvi	$(\text{nir} / \text{green}) - 1$	Mediana
nuaci	$UNTL * (1 - \sqrt{(\text{NDWI} - a\text{NDWI})^2 + (\text{NDVI} - a\text{NDVI})^2} + (\text{NDBI} - a\text{NDBI})^2)$	Mediana
pri	$(\text{blue} - \text{green}) / (\text{blue} + \text{green})$	Mediana

### 4. Clasificación

Para generar la clasificación, se desarrolló un algoritmo de procesamiento en seis etapas: cuatro de ellas dedicadas al proceso de clasificación y las dos restantes a la etapa de post-clasificación. Los resultados fueron generados y almacenados en Google Earth Engine. El proceso incluyó la definición de las áreas mineras, la recolección de muestras y la clasificación, seguido de la integración de todas las zonas mineras identificadas en una única capa. Finalmente, esta capa se integró con el mapa general de la colección 6 de MapBiomias Amazonía. A continuación, se describen en detalle cada uno de los pasos asociados a la clasificación.

#### 4.1. Búsqueda y actualización de datos

Se lleva a cabo una búsqueda exhaustiva de fuentes de datos auxiliares (mapas, reportes, informes) con el objetivo de identificar las principales áreas mineras. A partir de esta información, se genera una tabla con las coordenadas de cada zona minera, que posteriormente es exportada a una herramienta de sistemas de

información geográfica (SIG) para su visualización espacial en forma de un mapa de puntos (ver Figura 2).

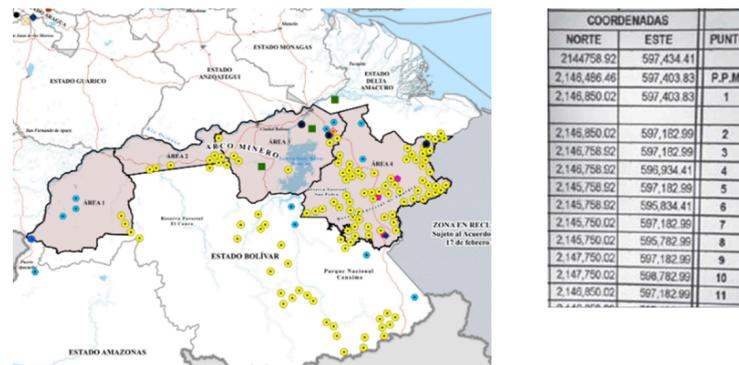


Figura 2: Tabla de coordenadas y mapa de puntos con la localización de las zonas mineras

#### 4.2. Preparación de los datos

El mapa de puntos de las zonas mineras se importa a Google Earth Engine como referencia espacial para delimitar las áreas mineras y recolectar muestras de entrenamiento. Como mapa base, se utilizan imágenes de alta resolución de Google Maps y los mosaicos Landsat (ver Sección 2). Con estos datos, se identifican y generan polígonos envolventes alrededor de las zonas mineras, abarcando una o varias minas (ver Figura 3). Estos polígonos se delimitan manualmente utilizando las herramientas de dibujo de Google Earth Engine, asegurando que las áreas mineras y no mineras fueran proporcionales para facilitar una clasificación precisa por parte del algoritmo. Para cumplir con este requisito, se dibuja un gran número de polígonos en ciertas regiones.

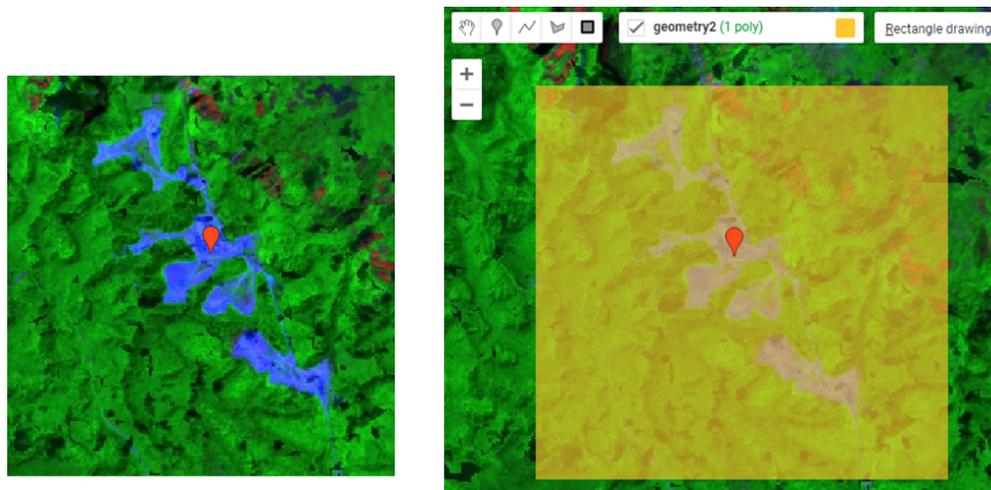


Figura 3: Polígono de zonas mineras

#### 4.3. Muestras de entrenamiento (minería/no minería)

Dentro de los polígonos que delimitan las áreas mineras de interés (Sección 4.2, ver Figura 2), se definen manualmente dos tipos de polígonos de entrenamiento: el primero, denominado 'Mina' (ID 30), que abarca exclusivamente el área de explotación

minera; y el segundo, denominado 'No\_Mina' (ver Figura 4), que incluye la clase 'no observado' o aquellas clases que no corresponden a minería (ID 27).

Las muestras de entrenamiento seleccionadas dentro del área de interés (polígonos envolventes) se utilizan para generar la clasificación de las zonas mineras. Todo lo que queda fuera de estas áreas delimitadas no es clasificado, ya que se considera como una zona no minera o asignada a otra clase dentro de la leyenda de MapBiomás.

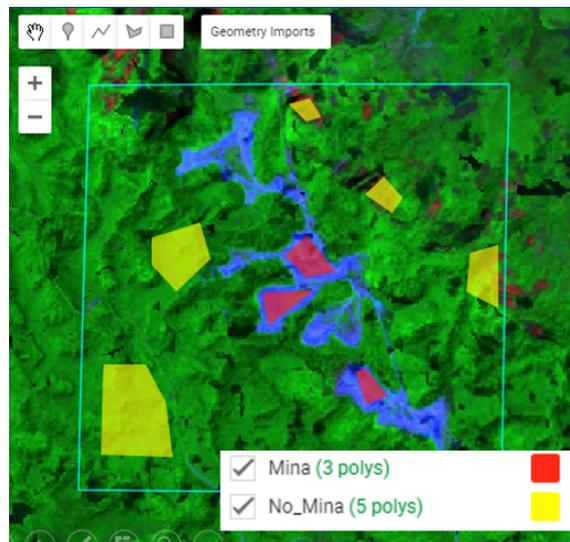


Figura 4: Polígono de Mina y No-Mina

#### 4.4. Clasificación con algoritmo Random Forests

La clasificación de las zonas mineras se realiza en la plataforma Google Earth Engine utilizando el algoritmo Random Forests (bosques aleatorios) (Breiman, 2001), basado en la metodología desarrollada por el equipo de MapBiomás Brasil (Solutions in Geoinformation, Solved, Brasil) para la identificación de minería industrial en ese país.

La clasificación se lleva a cabo dentro de las áreas delimitadas por los polígonos dibujados, tal como se describe en la Sección 4.2 de este documento. El resultado es una imagen compuesta por píxeles con valores binarios que indican la presencia o ausencia de zonas mineras (el área definida se muestra en color rojo en la Figura 5), dentro de cada una de las cuadrículas previamente establecidas.



Figura 5: Der. Mina de bauxita. Izq. Clasificación de zona minera

## 5. Post-clasificación

Los resultados de la clasificación son sometidos a una serie de filtros con el objetivo de reducir inconsistencias temporales, vacíos de información y otros problemas inherentes al proceso de clasificación. Para este propósito, se aplican un filtro espacial y un filtro de frecuencia.

### 5.1 Filtro espacial

El filtro espacial se basa en la función "connectedPixelCount" de Google Earth Engine, que identifica conjuntos de píxeles vecinos que comparten el mismo valor. Al aplicar esta función y establecer una cantidad mínima predefinida de píxeles idénticos, aquellos que no tienen conexión con otros (es decir, que presentan valores diferentes) se clasifican como píxeles aislados y son reclasificados (ver Figura 6). Para ambas clases (ID: 27 y ID: 30), la conexión mínima establecida es de 5 píxeles, lo que equivale aproximadamente a 0.5 hectáreas.

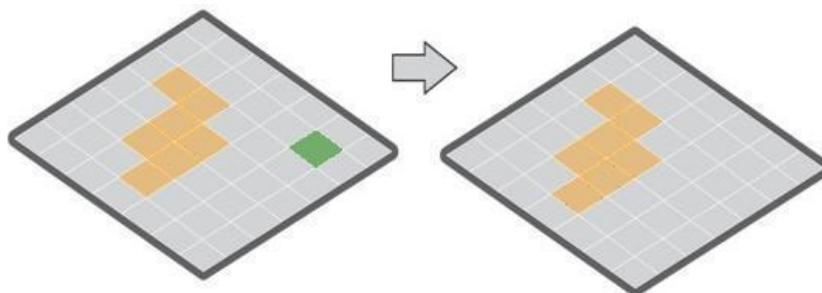


Figura 6: Esquema de la funcionalidad del filtro espacial

### 5.2 Filtro de frecuencia

Este filtro considera la frecuencia de ocurrencia de las clases minería (ID: 30) y no observado (ID: 27) a lo largo de toda la serie temporal. Su función es rellenar o

reemplazar el valor de la clase en aquellos años donde no se alcanza el valor mínimo de frecuencia, asignando el valor correspondiente de la clase predominante.

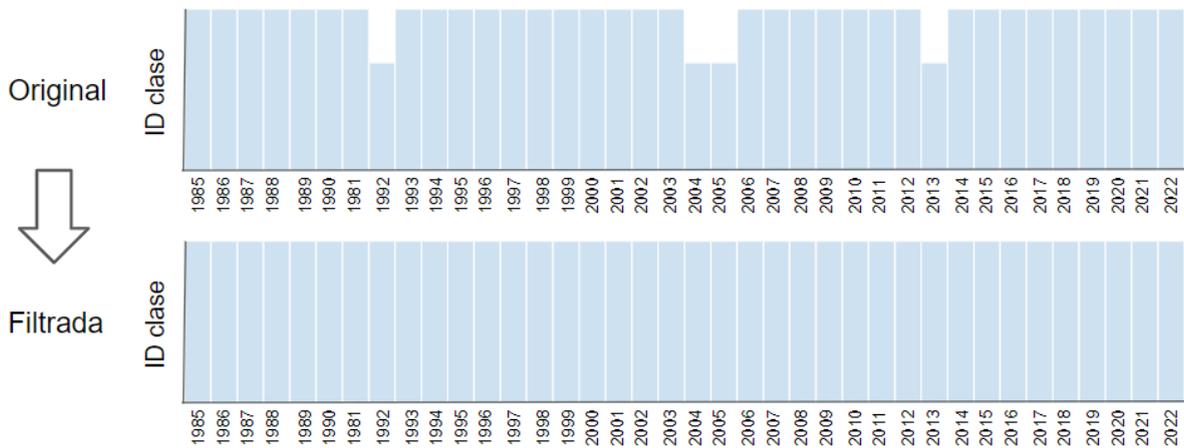


Figura 7: Efecto del filtro de frecuencia

### 5.3 Integración

Al obtener la clasificación de las zonas mineras se hace una primera integración de todas las zonas identificadas con minas para cada país (figura 8), y este resultado se integra al mapa general de clasificación nacional (figura 9). Esta integración se genera con base a una serie de reglas de prevalencia u orden de integración por jerarquía, para cada una de las clases dentro de la leyenda MapBiomás.

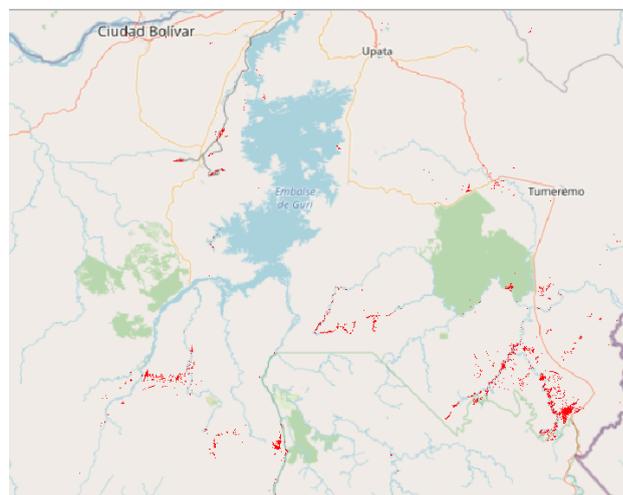


Figura 8: Clasificación de zona minera integrada



*Figura 9: Imágenes de referencia (resultado de la integración del tema transversal de minería en el mapa general, imagen de alta resolución de Google Earth y paisaje)*

## 6. Resumen metodológico

Los 6 pasos de clasificación del tema transversal de minería se representan esquemáticamente en el siguiente gráfico (figura 10).



Figura 10: Pasos para la clasificación de minería

## 7. Referencias

- 1) Breiman, Leo. "Random Forests." Machine Learning vol 45 (2001): 5–32.
- 2) Hartshorn, Scott. Machine Learning With Random Forests And Decision Trees A Visual Guide For Beginners. Amazon.com Services LLC, 2016.