

# Clasificación de coberturas en inmediaciones de la Sierra Nevada de Santa Marta a través de machine learning e imágenes Landsat

Juan Pablo Bedoya-Muñoz <sup>a</sup>

<sup>a</sup> Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia. [jupbedoyamu@unal.edu.co](mailto:jupbedoyamu@unal.edu.co)

Received: December 09th, 2022. Received in revised form: December 09th, 2022.

## Abstract

A multiclass classification problem was developed in which the landcovers around the Sierra Nevada de Santa Marta are defined (SNSM), for this purpose information obtained from Landsat 09 C2 satellite images dated 05/23/2022 in its different bands and on which the normalized difference indices are calculated : NDSI, NDGI, NDSII, NWI, NVDI. In addition to a Digital Elevation Model, the following covers are classified on these independent variables: vegetation, clouds, glacier, bare soil, lakes, and shadows. Different supervised and assembled methods are applied, which present a better fit to the data sets with which the models learn, with metrics of 95%, 98% and 100% in Linear Discriminant Analysis, K-nearest neighbors, and Random Forest (LDA, KNN and RF).

**Keywords:** classification, landcover, Sierra Nevada de Santa Marta, machine learning, supervised models, LDA, KNN, RF.

## Resumen

Se desarrolló un problema de clasificación multiclase en la cual se definen las coberturas de tierra alrededor de la Sierra Nevada de Santa Marta (SNSM), para ello se utilizó información obtenida de imágenes satelitales Landsat 09 C2 con fecha 23/05/2022 en sus diferentes bandas y sobre la cual se calculan los índices de diferencias normalizadas: NDSI, NDGI, NDSII, NWI, NVDI, además de un Modelo de Elevación Digital, sobre estas variables independientes se clasifican las siguientes coberturas: vegetación, nubes, glaciar, suelo desnudo, lagos, y sombras. Se aplican diferentes métodos supervisados y ensamblados, los cuales presentan un mejor ajuste a los conjuntos de datos con los cuales se les enseña a los modelos, con métricas del 95%, 98% y 100% en Análisis Discriminante Lineal, K-vecinos más cercanos, y Bosques aleatorios (LDA, KNN y RF, respectivamente por sus siglas en inglés).

**Palabras clave:** clasificación, coberturas; Sierra Nevada de Santa Marta, machine learning, modelos supervisados, LDA, KNN, RF.

## 1. Introducción

El estudio de los glaciares tropicales de manera directa o en campo es una labor compleja puesto que estos se encuentran en las cumbres de las montañas andinas, de esta manera se conocen como glaciares de montaña (IDEAM, 2000), por lo que su estudio podría ser lento y costoso, inicialmente estas coberturas podían ser delimitadas usando fotografías tomadas desde la superficie o fotografías aéreas pero poseían los mismos inconvenientes en cuanto a eficiencia, posterior a esto se desarrollaron técnicas de sensores remotos, los cuales trabajan usando los registros de reflectancia de las diferentes coberturas a determinadas bandas espectrales, por ejemplo imágenes satelitales de la misión Landsat, los cuales son datos abiertos y de fácil acceso, aprovechando esto es posible diferenciar coberturas a partir de estas técnicas, incluso se han desarrollado índices espectrales para resaltar algunas en específico, por ejemplo, el índice de diferencias normalizadas de nieve (NDSI, por sus siglas en inglés) permite la separación de zonas cubiertas por nieve de nubes u otras coberturas, sin embargo, no es un método automático y preciso, por lo tanto realizar composiciones en falso color como RGB-743 para Landsat 9 resulta conveniente para la delimitación de cobertura glaciar (Molano, 2022).

Sin embargo, estas técnicas siguen teniendo problemas en algunos casos, como en la detección de superficies glaciares cubiertas de escombros o rocas, llamados en ocasiones glaciares negros, generando la búsqueda de nuevas formas para extraer y delimitar de manera más precisa las superficies glaciares, y coberturas anexas a estos, es así como el aprendizaje de automático de máquinas (machine learning, ML, por sus siglas en inglés) se ha convertido en una herramienta muy usada, con muy buenos resultados en términos de precisión, en diferentes ámbitos incluyendo la extracción de coberturas, esto genera mejores resultados en cuanto a la precisión y eficiencia en términos operacionales que finalmente sirve para estimar entre otras cosas, balances de masa o tasas de retroceso glaciar (Manoj, 2011) y (Ambinakudige, 2022).

Los modelos usados para la identificación de superficies glaciares abarcan tanto modelos no supervisados, como supervisados y ensamblados. Usando como variables predictoras los diferentes índices espectrales y modelos de elevación digital. “Support Vector Machine” (SVM), Redes Neuronales Artificiales (ANN), y “Random Forest” (RF), fueron usados en la clasificación de glaciares en la cuenca de Hunza en Pakistán (Khan, 2020), y así mismo modelos como “K-Nearest Neighbors” (KNN), “Gradient Boosting” (GB), Árboles de decisiones (DT), “Multi-Layer

Perceptron” (MLPC) han sido usados con coeficientes kappa por encima de 0.8, este último valor en el peor de los casos. (Ambinakudige, 2022).

## 2. Métodos, datos y procesamiento

Para la clasificación de coberturas se utilizará una imagen satelital Landsat 9 con sensor OLI-2 para las bandas del 1 al 9, la cual consta de una resolución espacial de 30m por pixel, y una resolución radiométrica de 8 bits, y resolución en la banda pancromática de 15m, con fecha del 23 de mayo de 2022. Tomada del Servicio Geológico de Estados Unidos (USGS, por sus siglas en inglés) a través de su explorador para la visualización de bases de datos sobre imágenes satelitales (earthexplorer.usgs.com), además de un modelo de elevación digital obtenidos de esta misma fuente con una resolución espacial similar.

El procesamiento consta de tres fases principales, las cuales son, preprocesamiento, procesamiento y posprocesamiento.

### 2.1 Preprocesamiento

Etapas en la cual se descarga la imagen satelital y se hacen las correcciones necesarias, bien sean atmosféricas, geométricas y recorte al área de interés. Pero, debido a que se toman imágenes con un nivel de corrección 2 y asociadas a la colección 2, es decir, imágenes preprocesadas por las misiones Landsat con todas las correcciones tanto atmosféricas como geométricas, las cuales se encuentran listas para realizar una etapa de procesamiento con estas (EROS, 2020), la única corrección que se hace necesaria es un recorte al área de interés, el cual corresponde a un rectángulo de 2845 km<sup>2</sup> que abarca las superficies glaciares de la Sierra Nevada de Santa Marta (Figura 1).

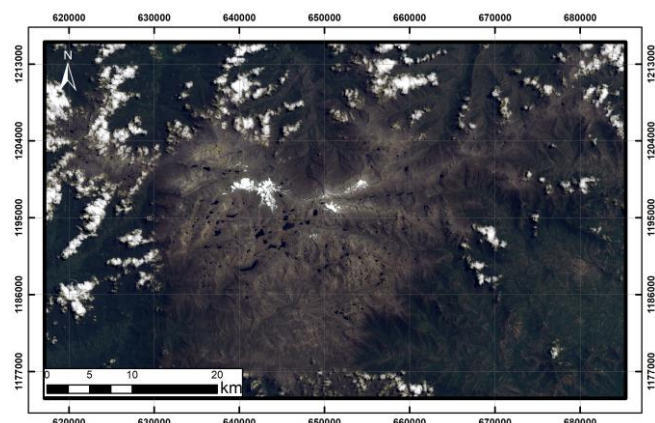


Figura 1. Mascara aplicada a la imagen satelital L9, área de interés. Fuente: Autor.

Posterior a este recorte, se calcularán los diferentes índices espectrales, usando ArcGIS v10.5 y a través de un álgebra de rasters, los cuales serán las variables predictoras en los modelos de machine learning a desarrollar, los cuales son: índice de nieve de diferencia normalizada (NDVI, por sus siglas en inglés) propuesto por primera vez por Dozier en 1989 (Dozier, 1989), índice de glaciar, nieve y hielo, y vegetación, todos de diferencia normalizada (NDGI, NDSII y NDVI respectivamente) propuestos en diferentes artículos (Maestro. I., 2004), (Manoj, 2011) y (Monterroso, 2018). Además del nuevo índice de agua (NWI) el cual utiliza las bandas del azul, infrarrojo cercano, infrarrojo de onda corta 1 y 2, propuesto en 2018 (Ding, 2018).

Tabla 1. Índices espectrales calculados en ArcGIS y usados como variables predictoras.

Fuente: Adaptada de (Monterroso, 2018)

Nombre del índice	Fórmula	Descripción
Índice de diferencia normalizada para Nieve	$NDSI = \frac{Green - SWIR}{Green + SWIR}$	Usado para la extracción de cobertura nevada
Índice de diferencia normalizada para Glaciar	$NDGI = \frac{NIR - Green}{NIR + Green}$	Usado para la extracción de cobertura nevada o glaciar
Índice de diferencia normalizada para Hielo y Nieve	$NDSII = \frac{Green - NIR}{Green + NIR}$	Usado para diferenciar entre hielo y nieve
Índice de diferencia normalizada para Vegetación	$NDVI = \frac{NIR - Red}{NIR + Red}$	Usado para la extracción de las coberturas vegetales y suelos desnudos
Nuevo índice de agua	$NWI = \frac{Blue - (NIR + SWIR1 + SWIR2)}{Blue + (NIR + SWIR1 + SWIR2)}$	Usado para la extracción de cuerpos de agua

Creación de grupos de control (Figura 2), los cuales son tomados de manera arbitraria sobre cada cobertura deseada, es decir, creación de una máscara para vegetación, nubes, glaciares, suelo desnudo, lagos y sombras, las cuales se buscan que sean heterogéneas en la composición en verdadero color, es decir, una cobertura que represente en general y de la mejor manera cada clase.

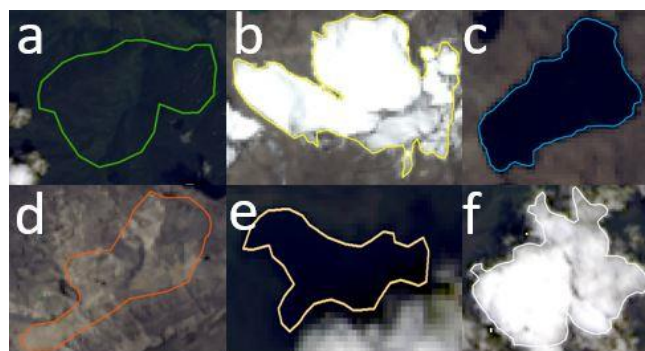


Figura 2. Grupos de control. a) vegetación b) glaciar c) lago d) suelo desnudo e) sombra f) nube. Fuente: Autor.

Posterior a esto, se hace una conversión a raster para generar los píxeles que abarca cada grupo de control y se genera una extracción de los valores de cada índice calculado y del DEM a través de la herramienta “sample” en ArcGIS, generando un archivo asociado a una Geodatabase, el cual es convertido a formato xls para posteriormente importarlo a Python para iniciar el procesamiento.

### 2.2 Procesamiento

Se procede a generar la lectura a través de Python de las diferentes tablas generadas por el cálculo de los índices, incluyendo además el Modelo de Elevación Digital, teniendo las variables predictoras junto a las clases a la cual corresponde cada entrada, se realiza un análisis exploratorio de datos, esto es concatenación de tablas, relleno de no datos en caso de encontrarse, a través de un

promedio entre la entrada anterior y posterior, grafico de cajas para entender las estadísticas de cada variable (Figura 3), correlación entre estas (Figura 4), con lo cual se busca entender las variables que mejor describen el modelo, y selección de varianza con VarianceThreshold.

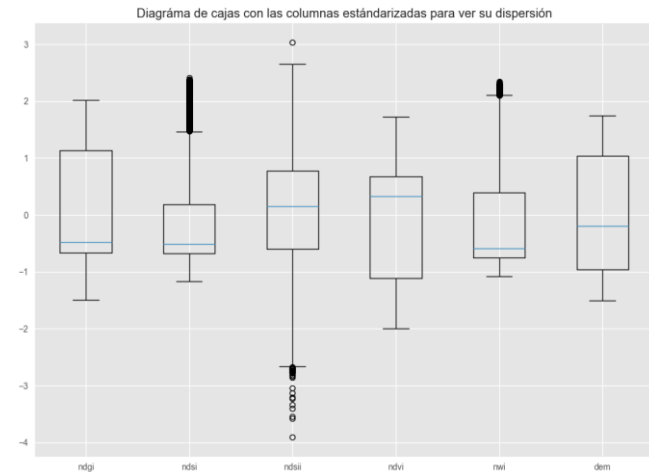


Figura 3. Diagrama de cajas para las variables predictoras.  
Fuente: Autor.

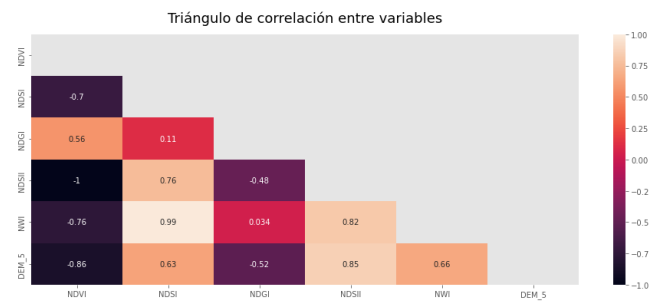


Figura 4. Triángulo de correlación entre variables.  
Fuente: Autor.

Seguido a esto, con el dataframe generado, se procede a realizar la partición de los datos en dos conjuntos, uno de entrenamiento que consta del 80% sobre el cual se procede a realizar la enseñanza de los diferentes modelos y otro del 20% restante de los datos como prueba para la predicción del modelo y medir su porcentaje de acierto.

	NDVI	NDSI	NDGI	NDSII	NWI	DEM_5	categoria
0	-0.044926	0.199722	0.002851	0.047771	-0.413084	4991	glaciar
1	-0.013605	0.204905	0.002855	0.016459	-0.398383	5007	glaciar
2	0.024920	0.006929	0.014650	-0.010274	-0.632121	5043	glaciar
3	0.048461	0.019522	0.002240	-0.046226	-0.562921	5055	glaciar
4	-0.031398	0.201222	-0.025102	0.006302	-0.406262	4987	glaciar
...	...	...	...	...	...	...	...
14025	0.301759	-0.091220	0.014150	-0.288842	-0.631418	2764	vegetacion
14026	0.303991	-0.092158	0.012830	-0.292301	-0.633051	2783	vegetacion
14027	0.313515	-0.098693	0.016565	-0.298500	-0.638920	2797	vegetacion
14028	0.303921	-0.103209	0.012703	-0.292347	-0.638347	2807	vegetacion
14029	0.310574	-0.109409	0.014644	-0.297282	-0.641550	2819	vegetacion

Figura 5. Dataframe generado para la particionar en los dos conjuntos.  
Fuente: Autor.

Finalmente, se procede a correr los diferentes modelos, los cuales corresponden a Análisis Discriminante Lineal (LDA), “K-Nearest Neighbors” (KNN), árbol de decisión (DT) y “Random Forest” (RF), a cada uno se le calcula su porcentaje de predicción, validación cruzada, y finalmente se procede a predecir cada pixel dentro de la máscara regional sobre la Sierra Nevada de Santa Marta y extraer su respectiva tabla.

	0	1	2	3	4	5
0	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
1	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
2	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
3	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
4	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
...	...	...	...	...	...	...
1386	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
1387	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
1388	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
1389	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion
1390	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion	vegetacion

Figura 6. Ejemplo de resultado, para Redes Neuronales Artificiales a conjunto de datos completo.  
Fuente: Autor.

Para el modelo de KNN se procede a buscar los hiperparámetros que pueden mejorar sus métricas a través de GridsearchCV.

## 2.2 Posprocesamiento

Una vez realizado la predicción de los diferentes modelos utilizados en este artículo se procede a la extracción de las tablas y graficado a través de la librería matplotlib. Además, la conversión a formato raster el cual será utilizado para visualizar en ArcGIS con una transparencia del 70% y comparar las diferentes clases calculadas con la imagen en verdadero color.

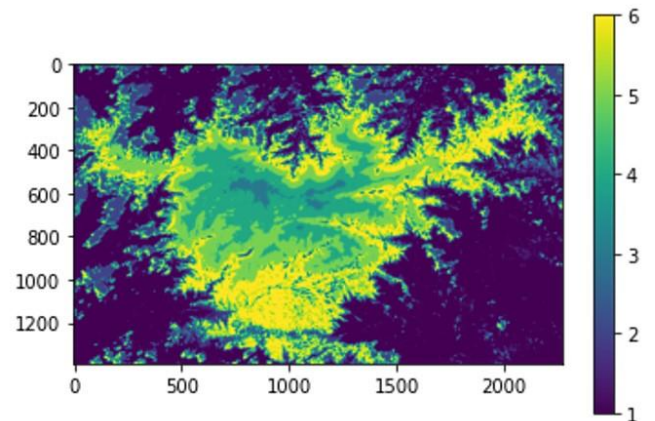


Figura 7. Ejemplo resultado predicción con Random Forest.  
Fuente: Autor.

## 3. Resultados y discusión.

### 3.1 Análisis exploratorio de datos

Se observa con el método de selección de variables a través de VarianceThreshold que la única variable que tiene una variación mayor al 75% (Figura 8), y que debería tomarse como representativa para los modelos sería el Modelo de Elevación Digital (DEM), pero



esto se explica debido a que las variables predictoras que se tomaron son índices de diferencias normalizadas, es decir son datos que no constan de unidades ni sentido físico y que además, presentan un rango entre -1 y 1, por lo tanto la varianza se disminuye y el selector nos entregara este resultado.

```
selector = VarianceThreshold(threshold=0.25)
selector.fit(X_train)
print(selector.get_support())
X_train.columns[selector.get_support()]

[False False False False False  True]
```

Figura 8. Selección de variables usando VarianceThreshold  
Fuente: Autor.

De acuerdo con esto, se siguen considerando todas las variables predictoras para el aprendizaje de los modelos, quedando con un conjunto de entrenamiento llamado de 11224 entradas y un conjunto de prueba llamado de 2806 entradas.

### 3.2 Análisis discriminante lineal (LDA)

Se toman 2 componentes sobre los cuales se trata de explicar la varianza de los datos. Se observa que, al entrenar el modelo con las 11224 entradas, y comparando los resultados que obtiene al hacer la predicción con las variables predictoras de prueba y comparando con las salidas de prueba que ya se conoce, se tiene una precisión de 98.07%, es decir el modelo aprende bien sobre estos datos, siendo castigado por la precisión en las sombras (Figura 9).

	precision	recall	f1-score	support
glaciar	1.00	0.97	0.99	511
lagos	0.93	0.99	0.96	129
nubes	1.00	0.94	0.97	657
sombra	0.70	0.99	0.82	79
suelos	0.98	1.00	0.99	432
vegetacion	1.00	1.00	1.00	998
accuracy			0.98	2806
macro avg	0.94	0.98	0.95	2806
weighted avg	0.98	0.98	0.98	2806

Figura 9. Métricas modelo LDA  
Fuente: Autor.

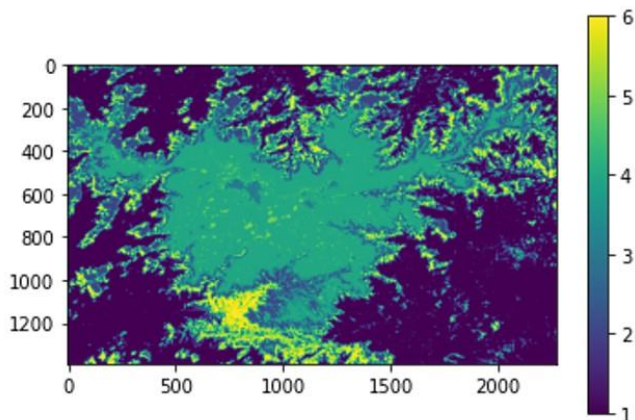


Figura 10. Resultado aplicación del modelo a conjunto de datos completo.  
1) Vegetación 2) Nubes 3) Glaciar 4) Suelo desnudo 5) Lagos 6) Sombras.  
Fuente: Autor.

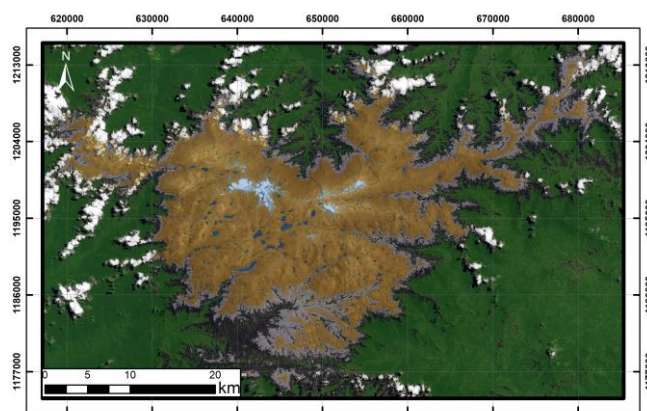


Figura 11. Transparencia del resultado del modelo con composición en verdadero color.  
Fuente: Autor.

Se observa que el modelo ajusta muy bien en cuanto a los verdaderos positivos en general, pero tiene también una alta tasa de falsos positivos, principalmente con respecto a las sombras y nubes. Sin embargo, la cobertura glaciar, queda muy bien delimitada, por lo que resulta muy efectivo en el momento de extraer coberturas glaciares únicamente.

### 3.3 "K-Nearest Neighbors" (KNN)

El hiperparámetro más importante para este modelo, es la cantidad de vecinos a tomar, para esto se procede a calcular a través de una curva de validación, en la cual se observa que el mejor valor de  $n$  es 2, debido a que posterior a esto y a medida que aumentan la cantidad de vecinos la precisión tanto para los datos de entrenamiento como en la validación cruzada va descendiendo, además de que se observa cómo va aumentando la varianza (Figura 12), con este valor la métrica de precisión es de 95.08%.

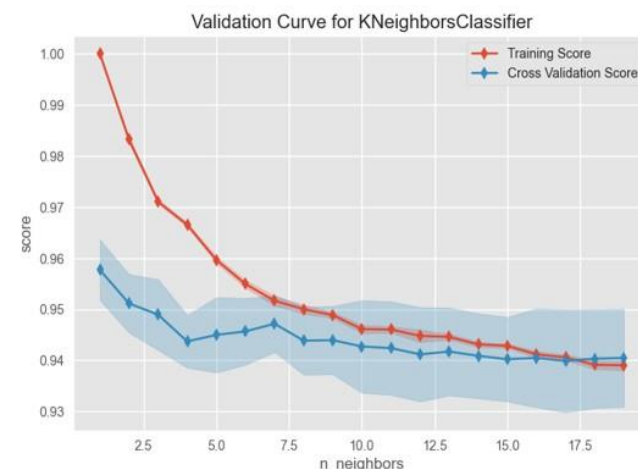


Figura 12. Curva de validación para el cálculo del mejor parámetro  $n$ .  
Fuente: Autor.

Además, se procede a generar un GridsearchCV para obtener los mejores hiperparámetros de tamaño de hoja,  $p$ , y  $n$ . para lo cual se obtienen que para todos el mejor valor posible es 1 y de esta manera se procede a correr el modelo según estos valores. Observando que las métricas aumentan en un punto porcentual, pasando a ser 96% de precisión, y en un 11% para la predicción de las sombras (Figura 13).

	precision	recall	f1-score	support
glaciar	1.00	1.00	1.00	511
lagos	1.00	1.00	1.00	129
nubes	0.91	0.97	0.94	657
sombra	0.32	0.23	0.26	79
suelos	1.00	1.00	1.00	432
vegetacion	0.97	0.94	0.95	998
accuracy			0.95	2806
macro avg	0.86	0.86	0.86	2806
weighted avg	0.95	0.95	0.95	2806

	precision	recall	f1-score	support
glaciar	1.00	1.00	1.00	511
lagos	1.00	1.00	1.00	129
nubes	0.96	0.93	0.95	657
sombra	0.43	0.35	0.39	79
suelos	1.00	1.00	1.00	432
vegetacion	0.96	0.98	0.97	998
accuracy			0.96	2806
macro avg	0.89	0.88	0.88	2806
weighted avg	0.96	0.96	0.96	2806

Figura 13. Métricas para KNN, con 2 vecinos, y con hiperparámetros encontrados con GridSearchCV.  
Fuente: Autor.

Finalmente se observa que las predicciones, para lagos, sombras, y suelos desnudos presenta una cantidad considerable de falsos positivos (Figura 14).

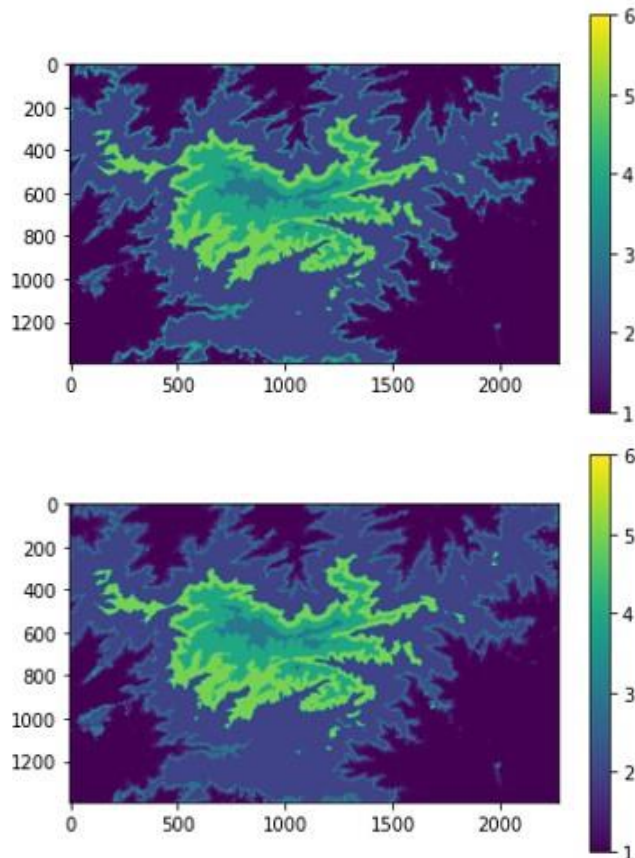


Figura 14. Resultado para KNN, con 2 vecinos, y con hiperparámetros dados por GridSearchCV.

1) Vegetación 2) Nubes 3) Glaciar 4) Suelo desnudo 5) Lagos 6) Sombras.  
Fuente: Autor.

### 3.4 “Decision Tree” (DT)

Al correr el modelo se obtiene un árbol de 45 hojas con una profundidad máxima de 12 ramas, las cuales generan una precisión de acierto muy alta, 99%, inclusive para acertar las sombras dentro del conjunto de prueba (Figura 15). Además, se observa como este, le da bastante peso a la variable DEM para empezar la clasificación (Figura 16), lo que significa que en el momento de tomar los grupos de control, es un hecho determinante la zona altitudinal sobre la que esta, lo que explica su distribución de falsos positivos de las sombras sobre el conjunto de datos completo (Figura 17).

	precision	recall	f1-score	support
glaciar	1.00	1.00	1.00	511
lagos	1.00	1.00	1.00	129
nubes	0.99	0.99	0.99	657
sombra	0.90	0.89	0.89	79
suelos	1.00	1.00	1.00	432
vegetacion	1.00	1.00	1.00	998
accuracy			0.99	2806
macro avg	0.98	0.98	0.98	2806
weighted avg	0.99	0.99	0.99	2806

Figura 15. Métricas para DT. 45 hojas, 12 profundidad máxima.  
Fuente: Autor.

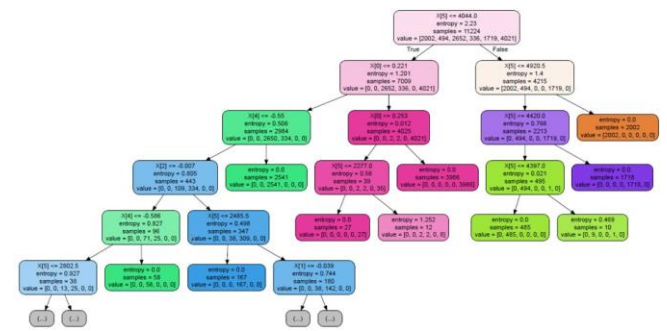


Figura 16. Árbol de decisión.  
Fuente: Autor.

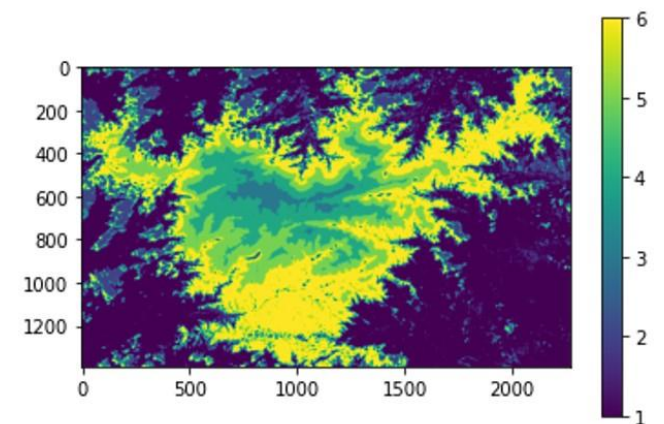


Figura 17. Resultados DT para conjunto de datos completo.  
1) Vegetación 2) Nubes 3) Glaciar 4) Suelo desnudo 5) Lagos 6) Sombras.  
Fuente: Autor.

### 3.5 “Random Forest” (RF)



Random Forest es el método con las métricas más altas para la clasificación de las diferentes coberturas (Figura 18), generando una precisión casi del 100% sobre todas las clases, incluyendo sombras, pero esto es sobre los datos de entrenamiento y de prueba, y se observa que el resultado sigue presentando una cantidad considerable de falsos positivos sobre las sombras y lagos (Figura 19). Sin embargo, es el segundo modelo después de LDA, que mejor delimita la cobertura glaciaria (Figura 20) por lo tanto es un buen modelo según los datos con los cuales se entrenó para extraer dicha cobertura para facilitar futuros análisis.

	precision	recall	f1-score	support
glaciar	1.00	1.00	1.00	511
lagos	1.00	1.00	1.00	129
nubes	1.00	0.99	1.00	657
sombra	0.95	0.96	0.96	79
suelos	1.00	1.00	1.00	432
vegetacion	1.00	1.00	1.00	998
accuracy			1.00	2806
macro avg	0.99	0.99	0.99	2806
weighted avg	1.00	1.00	1.00	2806

Figura 18. Métricas para RF.  
Fuente: Autor.

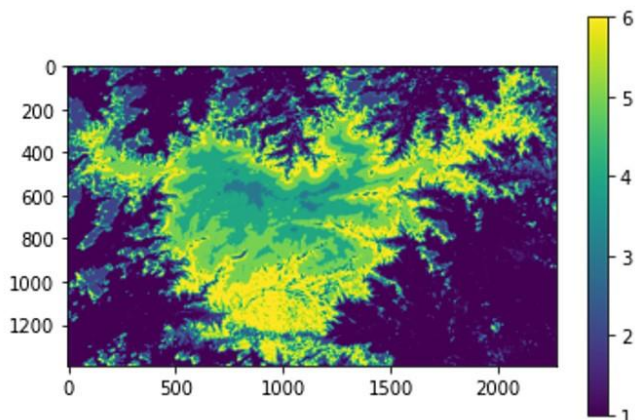


Figura 19. Resultados RF para conjunto de datos completo. 1) Vegetación 2) Nubes 3) Glaciar 4) Suelo desnudo 5) Lagos 6) Sombras.  
Fuente: Autor.

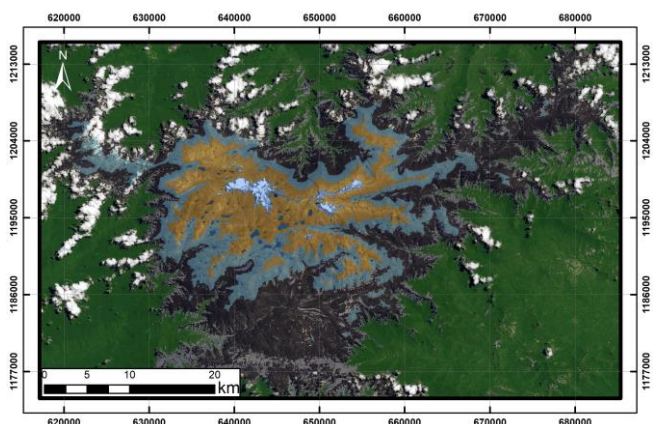


Figura 20. Resultado RF comparado a composición verdadero color.  
Fuente: Autor.

#### 4. Análisis y conclusiones.

Si bien se trabajan modelos basados en datos, es importante considerar el sentido físico del problema, puesto que los resultados muestran que se deben descartar variables solo por presentar variación menor al 75% y las cuales se deben entender como índices de diferencias normalizadas por lo tanto no resulta correcto establecer el grado de importancia dentro de estos modelos solo con selección de variables.

La Tabla 2 resume las métricas obtenidas por los diferentes modelos empleados en la delimitación o clasificación de las coberturas, en general todos presentan buenos aprendizajes y capacidad predictiva, y sin embargo, se observa que pese a que son "Random Forest" y "Decision Tree" los que presentan los mejores valores, es decir los métodos ensamblados, respecto a la precisión de los datos con sus conjuntos de prueba, esto tal vez porque son muy sensibles a los grupos de control que se escoja, en términos de que posiblemente le dan mucho peso a la variable DEM para clasificar, y tal vez por esto, se entrenan para predecir una cantidad impresionante de sombras, cuando estas por sentido físico del problema no dependen de esta. Además, el mejor modelo en términos de que genera menos falsos positivos dentro de las clases es Análisis Discriminante Lineal (LDA).

Tabla 2. Resumen métricas para los diferentes modelos empleados en el presente artículo.  
Fuente: Autor.

Modelo	Precisión
LDA	98.07%
KNN	95.08%
DT	99%
RF	100%

Es importante considerar entrenamientos con una base de datos mayor, puesto que una de las posibles respuestas frente a la aparición de altas cantidades de falsos positivos es el desbalance de 14.030 datos frente a 3.164.525 datos como el conjunto completo para correr los modelos ya entrenados, estamos en un escenario de underfitting, pese a tomar clases de control representativas teniendo en cuenta el conocimiento físico del problema.

Los modelos presentan diferencias significativas respecto a sus predicciones, es decir frente a sus aprendizajes del modelo basado en datos, los cuales son evidenciados en la comparación de los modelos frente a frente, la cual se observa en la siguiente Tabla 3.

Tabla 3. Comparación resultados entre modelos.  
Fuente: Autor.

Modelos	Comparación	Modelos	Comparación
LDA - KNN	0.56	KNN - RF	0.64
LDA - RF	0.68	KNN - DT	0.58
LDA - DT	0.72	DT - RF	0.91

#### Referencias

- [1] Ambinakudige, S. I. (2022). Estimation of area and volume change in the glaciers of the Columbia icefield, Canada using machine learning algorithms and Landsat images. *Remote sensing applications: Society and Environment*, 26, 1-11.
- [2] Ding, J. C. (2018). Monthly and annual temperature extremes and their changes on the Tibetan Plateau and its surroundings during 1963-2015. *Nature*, 1-23.
- [3] Dozier, J. (1989). Spectral Signature of Alpine Snow Cover from the Landsat Thematic Mapper. *Remote Sensing of Environment*, 9-22.
- [4] EROS. (2020). Landsat 8-9 OLI/TIRS Collection 2 Level-2 Science Products.

- [5] IDEAM. (2000). Los glaciares Colombianos, expresión del cambio climático. *Ministerio de Medio Ambiente*, 1-19.
- [6] Khan, A. J. (2020). Machine-learning algorithms for mapping debris-covered glaciers: the hunza basin case study. *IEEE*, 8, 12725–12734.
- [7] Maestro, I., R. J. (2004). Análisis comparativo de distintos métodos de estimación de la superficie glaciar en el Pirineo axial aragonés a partir de imágenes Landsat. *GeoFocus*, 4., 79-92.
- [8] Manoj, K. A. (2011). Digital imagen information extraction techniques for snow cover mapping from remote sensing data. *Encyclopedia of Snow, Ice and Glaciers*, 213-232.
- [9] Molano, S. C.-N. (2022). Evaluación del retroceso glaciar de la Sierra Nevada del Cocuy, Colombia a partir de la clasificación de imágenes multisensor. *Boletín de Geología*, 44(1), 49-73.
- [10] Monterroso, M. L. (2018). Estimación del retroceso glaciar en los volcanes Nevado del Ruiz, Tolima y Santa Isabel, Colombia a través de imágenes ópticas y Din-SAR. *DYNA*, 85(206), 329-337.

**J.P. Bedoya-Muñoz**, estudiante de ingeniería geológica en Universidad Nacional de Colombia, Sede Medellín.