

## Data Dictionary

Feature	Description
<b>NDSI</b>	Normalized Difference Snow Index: $(\text{Green} - \text{SWIR}) / (\text{Green} + \text{SWIR})$
<b>NDVI</b>	Normalized Difference Vegetation Index: $(\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red})$
<b>SR_B2 to SR_B7</b>	Surface reflectance in Blue, Green, Red, NIR, SWIR1, SWIR2
<b>ST_B10</b>	Thermal infrared band
<b>QA_PIXEL</b>	Pixel quality indicator
<b>QA_RADSAT</b>	Radiometric saturation flag
<b>SR_QA_AEROSOL</b>	Aerosol quality from surface reflectance
<b>ST_ATRAN</b>	Atmospheric transmittance
<b>ST_CDIST</b>	Distance to cloud
<b>ST_DRAD, ST_EMIS, etc.</b>	Additional surface temperature quality metrics
<b>class</b>	Binary label: 1 = glacier, 0 = non-glacier

## Model Visualizations

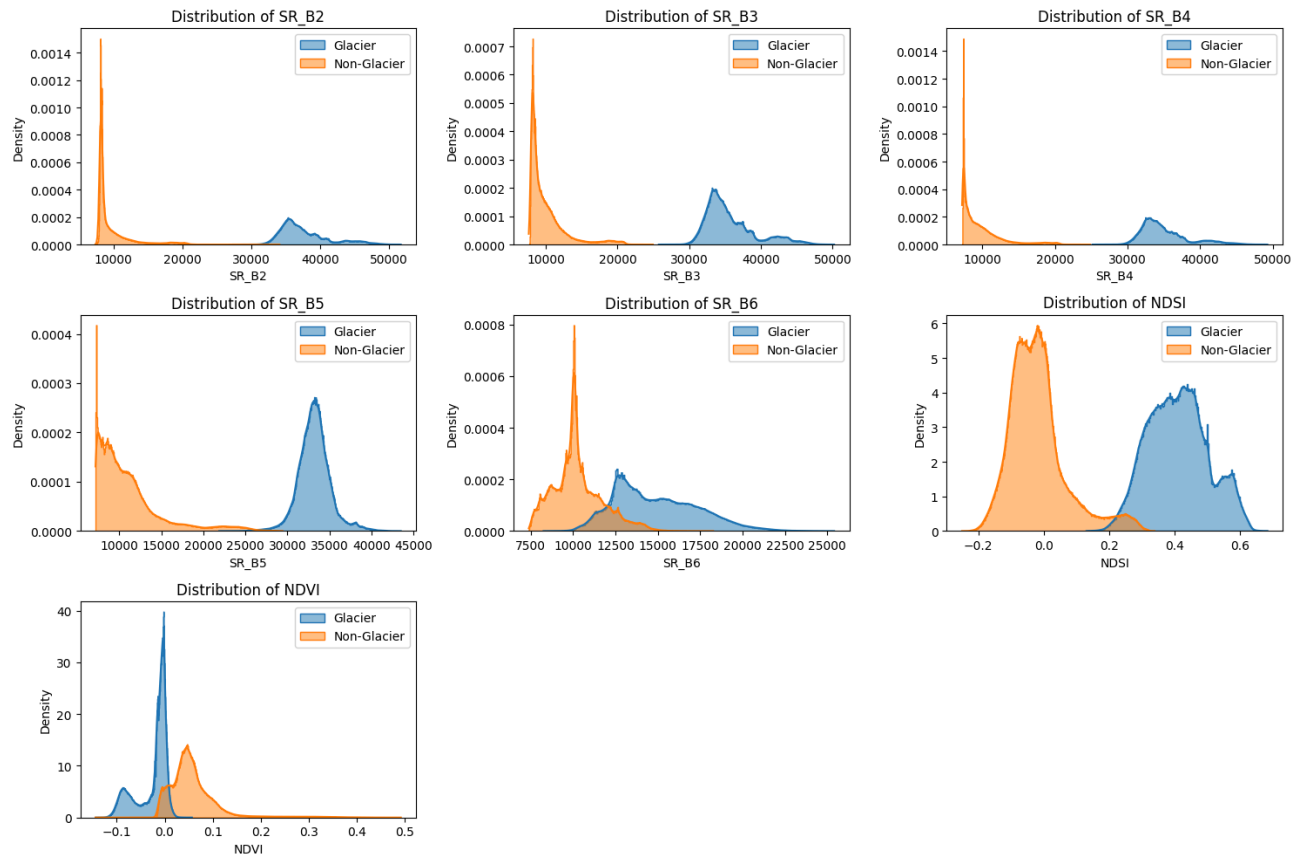
To better understand how the spectral bands and indices contribute to glacier classification, we visualized the distributions of key features across glacier and non-glacier classes using histograms, violin plots, density plots, and correlation analysis.

### Band Distributions

We plotted the distributions of SR\_B2–SR\_B6, NDVI, and NDSI across both classes:

- **Histograms** and **density plots** show clear separations between glacier and non-glacier pixels across most bands.

*Figure I: Histograms of Distribution of Each Spectral Band*

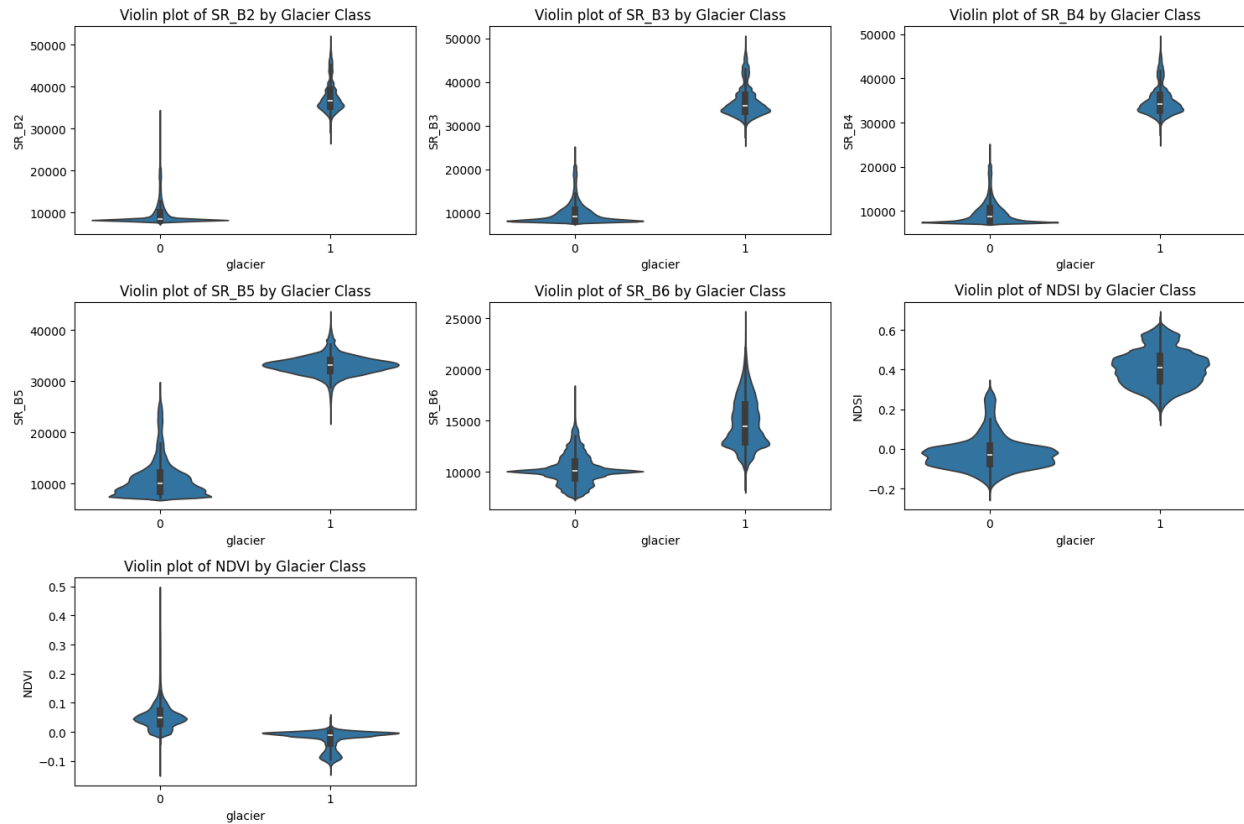


- **Glacier pixels** tend to have higher reflectance in bands SR\_B2–B5 and higher NDSI values.
- **Non-glacier pixels** are more dispersed but tend to cluster around lower reflectance and NDSI values.

## Violin Plots by Class

Violin plots further illustrate this separation:

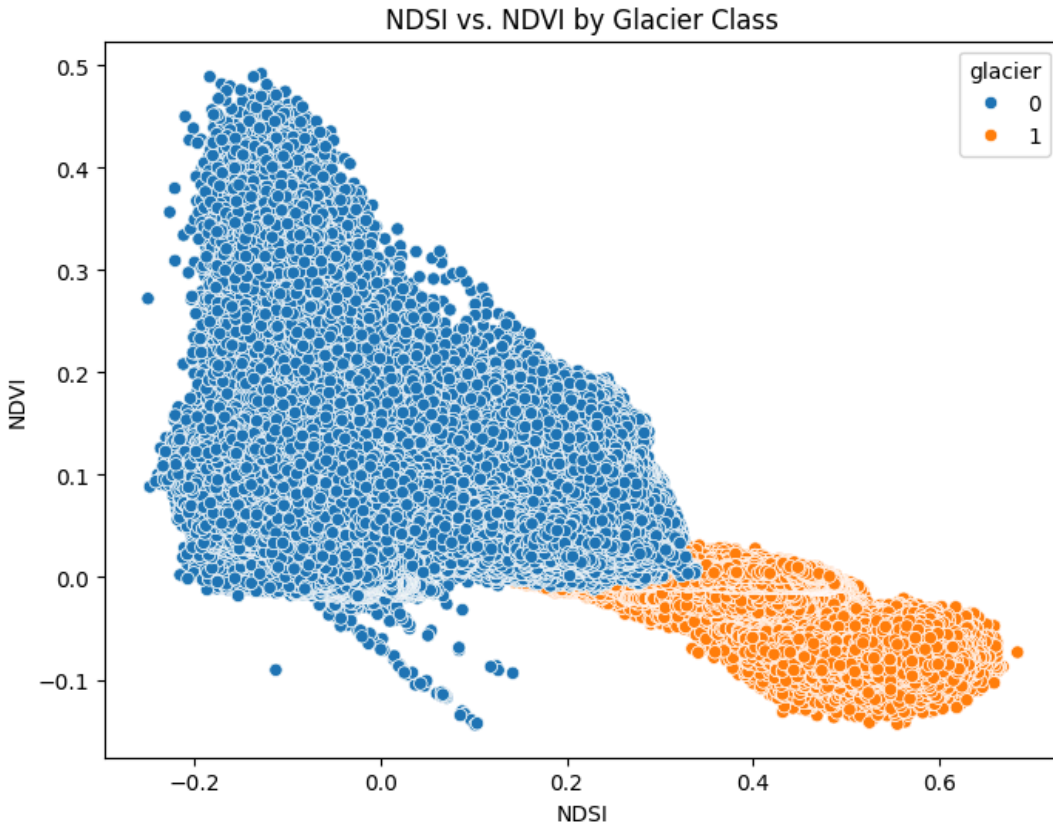
*Figure II: Violin Plots of Each Spectral Band by Glacier Class*



- Bands like **SR\_B5** and **NDSI** show tight, distinct clusters by class.
- **NDVI** exhibits a more subtle but still meaningful shift between glacier and non-glacier regions, with glaciers generally showing lower vegetation index values.

## NDSI vs. NDVI Scatterplot

Figure III: Scatterplot of NDSI vs. NDVI colored by glacier class



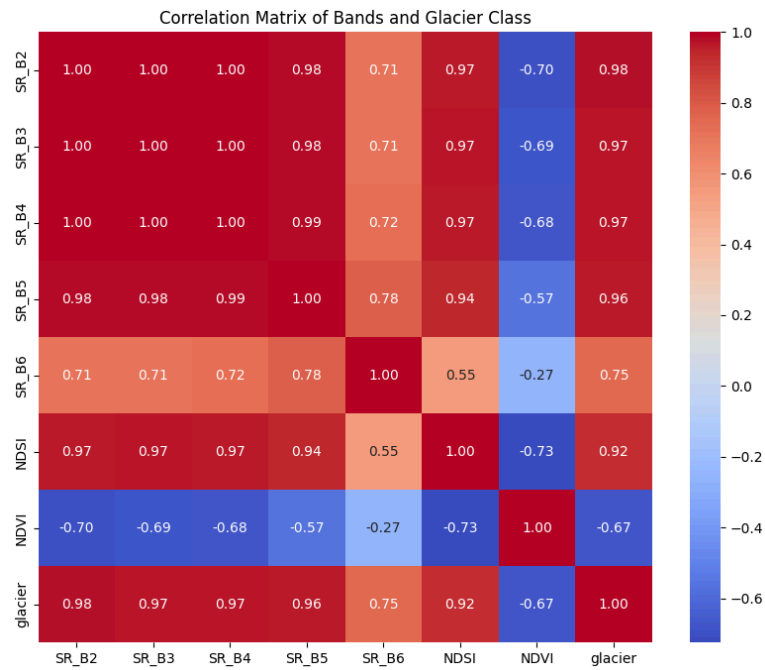
- Glacier pixels cluster in a region with **high NDSI and low NDVI**.
- Non-glacier pixels spread more broadly, with **higher NDVI and lower NDSI**, confirming their potential as a decision boundary in classification tasks.

## Correlation Matrix

The correlation heatmap shows strong positive correlations between the glacier class and:

- SR\_B2–SR\_B5 (~0.96–0.98)
- NDSI (0.92)

Figure IV: Correlation Matrix of Bands and Glacier Class



There's also a **negative correlation with NDVI (-0.67)**, reinforcing that glacier-covered areas are spectrally and vegetatively distinct from surrounding land cover.