# **SIIM-ISIC Melanoma Classification**

## Introduction:

Skin cancer is the most common type of cancer. It occurs due to the abnormal growth of skin cells, usually on the areas exposed to sunlight. Melanoma is responsible for 75% of skin cancer deaths, despite being the least common skin cancer. The American Cancer Society estimates over 100,000 new melanoma cases will be diagnosed in 2020. As with other cancers, early and accurate detection-potentially aided by data science-could make treatment more effective. Melanoma is a deadly disease, but if caught early, most melanomas can be cured with minor surgery.

Image analysis tools that automate the diagnosis of melanoma will improve dermatologists diagnostic accuracy. Better detection of melanoma has the opportunity to positively impact millions of people. We'll identify Melanoma in images of skin lesions. In particular, we’ll use images within the same patient and determine which are likely to represent a melanoma. Using patient-level contextual information may help the development of image analysis tools, which could better support clinical dermatologists.

## Goals:

The objective of the project is to accurately identify the likeliness that images of skin lesions of patients represent melanoma based on the data provided and certain features present in data

We are going to use different algorithms to implement SIIM-ISIC Melanoma Classification and check which method gives us the maximum model performance.

In short, we are trying to analyze the following points

1. Is it possible to predict if a person has cancer based on the data which is provided?
2. Understand which algorithm has the highest performance score
3. To apply concepts of transfer learning for improving the performance

## Data

* The images are provided in DICOM format. This can be accessed using commonly available libraries like pydicom, and contains both image and metadata.
* Images are also provided in JPEG.
* Metadata is also provided in CSV files.

## Columns:

* **image\_name** - unique identifier of the image
* **patient\_id** - unique identifier of the patient
* **sex** - the sex of the patient
* **age\_approx** - approximate age of the patient
* **anatom\_site\_general\_challenge** - location of imaged site
* **diagnosis** - information about the diagnosis
* **benign\_malignant** - indicator if the image is malignant or not (string)
* **target** - indicator if the image is malignant or not (binary)

## Importance:

* Image analysis tools that automate the diagnosis of melanoma will improve dermatologists diagnostic accuracy.
* Better detection of melanoma has the opportunity to positively impact millions of people.
* Early and accurate results can help in better providing better treatments

## Task Breakdown:

|  |  |
| --- | --- |
| **Task** | **Done by** |
| XGBoost | Rishvita |
| Logistic Regression | GnanaTej |
| CNN | Pranathi |
| Transfer Learning(VGG16) | Rishvita |
| Transfer Learning(DenseNet) | GnanaTej |
| Model Comparison | Pranathi |