**Task**: Detection of the presence of St. George in an image.

**Dataset:** [[](https://aerialogy-my.sharepoint.com/:f:/g/personal/rajeshwarie_chatterjee_starinco_com/EmkOcQDsmDRHizYJ9IyIY48BI1QG79nUBRHTo1uw2TxoDw?e=6EH8UN)Data Engineer Test](https://aerialogy-my.sharepoint.com/:f:/g/personal/rajeshwarie_chatterjee_starinco_com/EmkOcQDsmDRHizYJ9IyIY48BI1QG79nUBRHTo1uw2TxoDw?e=6EH8UN)

The two files, viz., georges.csv and non-georges.csv, contain the list of pictures with St. George and without St. George, respectively.

**Assignment:** Write code for training an appropriate model for the given task and explain your choice of test metrices and the corresponding training results. Also write code that takes an image as input and classifies it based on the presence and absence of St. George using the model trained earlier. Discuss the rationale behind the selection of your training model and possible improvements based on your classification results.

**Time to complete the project:** You can take anywhere between a few hours to a maximum of 5 days depending on the level of details and the number of approaches you'd like to try.

**Submission process:** Create a public repository on Github or Gitlab or Bitbucket with a custom manager for managing dependencies and a README with detailed comments on running and reproducing the project, if necessary. Alternatively, you can develop the project in Google Colab and share the project. Add relevant comments to your code elucidating different parts of the code base.

**Guidelines:** You can use every trick from your repertoire to solve the problem. You can take ready-made models, download additional data to get better results. In such a case, provide information to reproduce these steps or provide (link to) the data itself in the README or in the comments. Also, explain your reasons behind such steps.

**Assessment criteria:**

\* Reproducibility of code/research/training

\* Code quality

\* The accuracy of the model developed