**[PROJECT TITLE IN CAPITAL LETTERS]**

A Project Report Submitted

in Partial Fulfilment of the Requirements for the Degree of

**[TYPE DEGREE NAME]**

in

**[TYPE DEPARTMENT NAME]**

*by*

**[Type your full name] (Roll No. [Type your roll no.])**



#### to

**DEPARTEMENT OF COMPUTER SCIENCE**

**INDIAN INSTITUTE OF INFORMATION TECHNOLOGY KOTTAYAM-686635, INDIA**

#### April 2023

**DECLARATION**

We, **Basanth**,Nancy,Susmera and Thomas hereby declare that, this report entitled **“Visual Narration :An Xception based image captioning model”** submitted to Indian Institute of Information Technology Kottayam towards partial requirement of **Master of Technology** is an original work carried out by us under the supervision of **Dr. Keshab Nath** and has not formed the basis for the award of any degree or diploma, in this or any other institution or university. We have sincerely tried to uphold the academic ethics and honesty. Whenever an external information or statement or result is used then, that have been duly acknowledged and cited.

Kottayam-686635 **[Your Name]**

April 2019

**CERTIFICATE**

# This is to certify that the work contained in this project report entitled “ Predictive analytics model for natural disasters ’ submitted Basanth,Nancy,Susmera and Thomas to Indian Institute of Information Technology Kot- tayam towards partial requirement of Master of Technology has been carried out by [him/her] under my supervision and that it has not been submitted elsewhere for the award of any degree.

Kottayam-686635 (Dr. Keshab Nath)

April 2019 Project Supervisor

**Contents**

Contents

[1. Introduction 6](#_Toc133596781)

[2 Data 6](#_Toc133596782)

[2.1 Preprocess the Data 6](#_Toc133596783)

[2.2 Splitting the Data 7](#_Toc133596784)

[3 Different Models 7](#_Toc133596785)

[3.1 Logistic Regression 7](#_Toc133596786)

[3.2 Random Forest Regression 7](#_Toc133596787)

[Model provided and accuracy score of 0.99051 for test and train set 7](#_Toc133596788)

[3.3 Gradient Boosting Regression 7](#_Toc133596789)

[Model provided and accuracy score of 0.85701 7](#_Toc133596790)

[3.4 SVM 7](#_Toc133596791)

[4 Comparison of Models 7](#_Toc133596792)

# Introduction

Acute natural disasters with origins in the atmosphere, biology, geophysics, hydrology, or oceanography pose a constant threat to our way of life and the environment, which is still incredibly unpredictable. These occasions, which are usually referred to as natural disasters, not only affect millions people globally each year, but they may also have a disproportionately large impact on some regions and populations, which widens the intercommunal capacity gaps.

Predicting the natural disasters can help to take better decisions on preventive and action ons required to reduce the impact.

4 different models were considered to find best suitable model to predict the hurricanes in

# Data

## Preprocess the Data

Dataset Contained different types of disaster data with around 40K records .Hurricanes was selected to the incident type that will be predicted.

* + 1. **Data Cleaning**

After the null checks unwanted columns were removed from the data set. Columns fy\_declared", "declaration\_date", "declaration\_title", "hash", "last\_refresh", "id", "month\_year", 'TEMPERATURE\_ID', 'DATE',  'Country', 'Fahrenheit', 'AverageTemperatureUncertaintyF', 'year\_y', 'month\_y'are removed.

* + 1. **One Hot Encoding**

One hot encoding was used to convert the incident\_type column into numeric values. STATE, declaration\_type , place\_code and designated\_area are also encoded to convert into numeric values.

## Splitting the Data

Data is split into independent variables in the variable X and Independent variable in y. After splitting data is normalized using Standanrd scaler transformation.

# Different Models

Introductory lines Logistic regression, Random Model and Gradient Boosting and SVM were the models selected to predict hurricanes.Random Forest and Gradient Boosting are generally selected when the data set is larger and above 20000 rows.

## Logistic Regression

Model provided and accuracy score of 0.8999

## Random Forest Regression

## Model provided and accuracy score of 0.99051 for test and train set

## Gradient Boosting Regression

## Model provided and accuracy score of 0.85701

## SVM

# Comparison of Models

# Chart, bar chart Description automatically generated