



# **Data Science Intern at Data Glacier**

**Project:** Hate Speech Detection using Transformers (Deep Learning)

**Week 7:** Deliverables

**Name:** Ishwarya Arulvel

**Email:** ish351994@gmail.com

**Country:** Canada

**Specialization:** Data Analyst

**Batch Code:** LISUM14

**Date:** 16 Nov 2022

**Submitted to:** Data Glacier

## Table of Contents:

|                            |   |
|----------------------------|---|
| 1. Project Plan .....      | 3 |
| 2. Problem Statement ..... | 3 |

## 1. Project Plan

| Weeks    | Date         | plan  |
|----------|--------------|---|
| Weeks 07 | Nov 16, 2022 | Problem Statement, Data Collection, Data Report |
| Weeks 08 | Nov 23, 2022 | Data Preprocessing                              |
| Weeks 09 | Nov 30, 2022 | Feature Extraction                              |
| Weeks 10 | Dec 7, 2022  | Building the Model                              |
| Weeks 11 | Dec 14, 2022 | Model Result Evaluation                         |
| Weeks 12 | Dec 21, 2022 | Flask Development + Heroku                      |
| Weeks 13 | Dec 28, 2022 | Final Submission (Report + Code + Presentation) |

## 2. Problem Statement

The term hate speech is understood as any type of verbal, written or behavioral communication that attacks or uses derogatory or discriminatory language against a person or group based on what they are, in other words, based on their religion, ethnicity, nationality, race, color, ancestry, sex or another identity factor. In this problem, we will take you through a hate speech detection model with Machine Learning and Python.

Hate Speech Detection is generally a task of sentiment classification. So, for training, a model that can classify hate speech from a certain piece of text can be achieved by training it on data that is generally used to classify sentiments. So, for the task of hate speech detection model, We will use Twitter tweets to identify tweets containing Hate speech.