



# Data Science Intern at Data Glacier

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

**Week 7:** Deliverables

**Name:** Amir Ali

**University:** Warsaw University of Technology

**Email:** [amirali.researcher@gmail.com](mailto:amirali.researcher@gmail.com)

**Country:** Poland

**Specialization:** Data Science

**Batch Code:** LISUM08

**Date:** 18 May 2022

**Submitted to:** Data Glacier

## Table of Contents:

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

## 1. Project Plan

<b>Weeks</b>	<b>Date</b>	<b>plan</b>
Weeks 07	May 18, 2022	Problem Statement, Data Collection, Data Report
Weeks 08	May 25, 2022	Data Preprocessing
Weeks 09	June 1, 2022	Feature Extraction
Weeks 10	June 8, 2022	Building the Model
Weeks 11	June 14, 2022	Model Result Evaluation
Weeks 12	June 21, 2022	Flask Development + Heroku
Weeks 13	June 30, 2022	Final Submission (Report + Code + Presentation)

## 2. Problem Statement

The term hate speech is understood as any type of verbal, written or behavioural 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, colour, 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 a data that is generally used to classify sentiments. So for the task of hate speech detection model, We will use the Twitter tweets to identify tweets containing Hate speech.