

# Covid-19 Vaccine Tweet Sentiment Analysis

AML Course Group Project

Ian Eustis, Sarah Hellmann, Kiriane Chikhaoui, Alexander Hildreth

# Agenda

1. Introduction to Team and Project Goal
2. Project Development
3. End-Product Analysis
4. End-Product Demonstration



# 1. **Introduction**

Introduction to Team and Project Goal

# Team Information

- Social Media Health Organization (SMHO)
  - Our company focuses on data scraped from Twitter's API
  - With this data we aim to better understand the general public's opinion on current health topics

# Project Goal

- To analyze and better understand the U.S. public sentiment of the Covid-19 Vaccines in real time
- To compare and contrast the public's sentiment towards the different vaccines throughout the early vaccine rollout in the United States



2.

# Project Development

Project Overview and ML Workflow

# Project Overview

- Our end-to-end Python application scrapes real-time twitter data and continuously runs a sentiment analysis. It labels the vaccine-related tweets with a high/low sentiment and visualizes the results on a web application
- We utilized the following AWS services:
  - A DynamoDB table to hold our data
  - 3 Lambda functions that:
    - Collect data
    - Pull the model and vectorizer files and apply prediction values
    - Summarize predictions for visualizations
  - An S3 bucket to hold our model and vectorizer, aggregate files, and Lambda triggers
  - \*\*Elastic Beanstalk to host our web application



## ML Workflow

### 1. Download Data

A **Lambda function** scrapes twitter API, downloads, and deposits unlabeled vaccine related tweets into an unlabeled **DynamoDB table**.

### 2. Apply Label

The deposit **triggers another Lambda function** which pulls the vectorizer and model from an **S3 bucket** and applies the sentiment label to the record in the DB.

### 3. Aggregate Findings

Multiple more **Lambda functions** scan the labeled **DynamoDB** table to aggregate files for data visualization

### 4. Deploy App

The final **application**, hosted on **Elastic Beanstalk\*\*** pulls the aggregated data and analyzes our predictions with **Plotly charts** to visualize meaningful insights.





3.

# End-Product Analysis

Effectiveness & Potential Improvements

# Effectiveness

- Our final sentiment analysis successfully performed in providing us with a better understanding of the varying public opinion throughout the pivotal vaccine rollout phase during Spring 2021
- The final product is:
  - Informative
  - Scalable
  - Accurate (85%)
  - Reusable for future analysis

# Potential Improvements

- Adding more analytical visualizations to acquire a deeper understanding of the vaccine sentiment
- Allowing for other languages and adding more vaccine types in the analysis as they become available
- Incorporation of twitter engagement metrics into analysis (like, retweet, reply, quotes)
- Including sentiment mapping based on user/tweet location



4.

# End-Product Demonstration

Web Application and AWS Demonstration

# Thanks!

Any questions?