Google Cloud Vertex AI Hackathon

JustBePractical

SocioSense

Combat hate speech and offensive content with our sentiment analysis AI, elevating communication and engagement on modern social media platforms for businesses and individuals.

Vertex AI

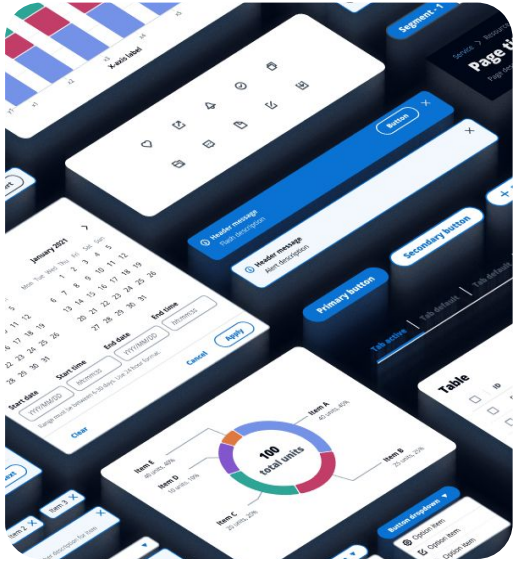
Google Cloud

lab lab ai

Overview

Introduction & Team	01
Our concept and aim	02
Sentiment Analysis AI	03
Transforming Online Presence Management	04
Navigating the Complexities of Online Presence	05
Building the Path to Success	06
Deployment and real world results	07
The future of the project	08

Introduction



Shuvam

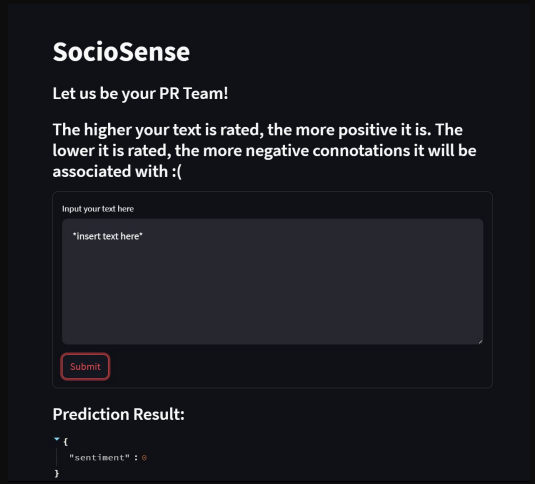
Developer

JustBePractical

Welcome to our team, JustBePractical, Today we want to share with you an intriguing concept that we believe has the potential to revolutionise the way businesses navigate the complex social media landscape, while also fostering a more inclusive and engaging online environment.. The integration of sentiment analysis into Vertex AI's AutoML Model allows for the automated identification and categorization of hate speech and offensive language across various social media platforms. By analyzing the sentiment behind users' words and phrases, this technology enables businesses and individuals to efficiently detect and respond to such content. This not only saves valuable time and resources but also acts as a mobile public relations team, helping individuals create content while minimizing discriminatory boundaries between online interactions.

Navigating the Complexities of Online Presence

In today's complex social media landscape, businesses and individuals face a significant challenge in effectively managing their online presence. Hate speech and offensive language on social media platforms not only tarnish reputations but also create a hostile environment that hinders effective communication and audience engagement. Navigating the digital battleground of public perception is increasingly difficult, with mass media disruptions and the detrimental effects of cancel culture lurking at every corner. It's crucial for businesses and individuals to break down discriminatory boundaries online and safeguard their reputation to thrive in the digital age.



How we achieved this:

1. Train the sentiment analysis from Google Vertex AI's Model Garden on our dataset.
2. Build a streamlit app with the help of the tutorial on lablabai's AI tutorials.
3. Create a front-end and back-end for the app.
4. Make use of the API to call the model to predict sentiment values of the entered text.

Problems

1. We had many problems with data cleaning and labelling some values for which the dataset was incomplete... overcame it by learning how to clean data using some pandas tutorials online.

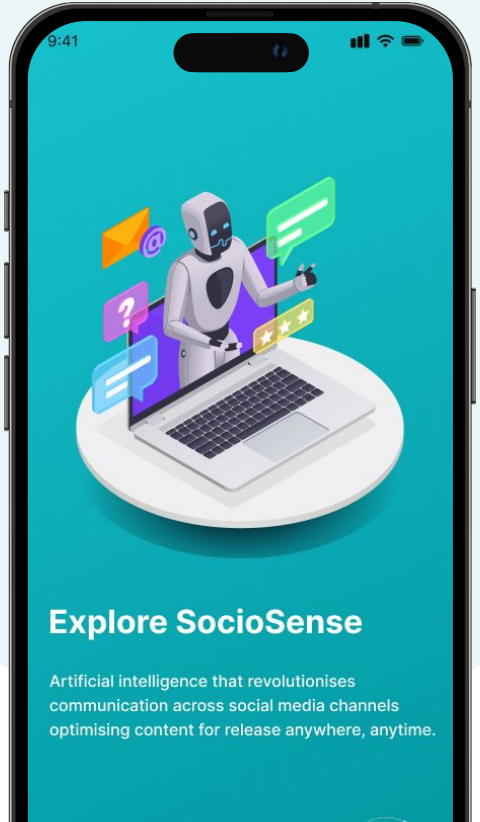
Neel Nair
Developer

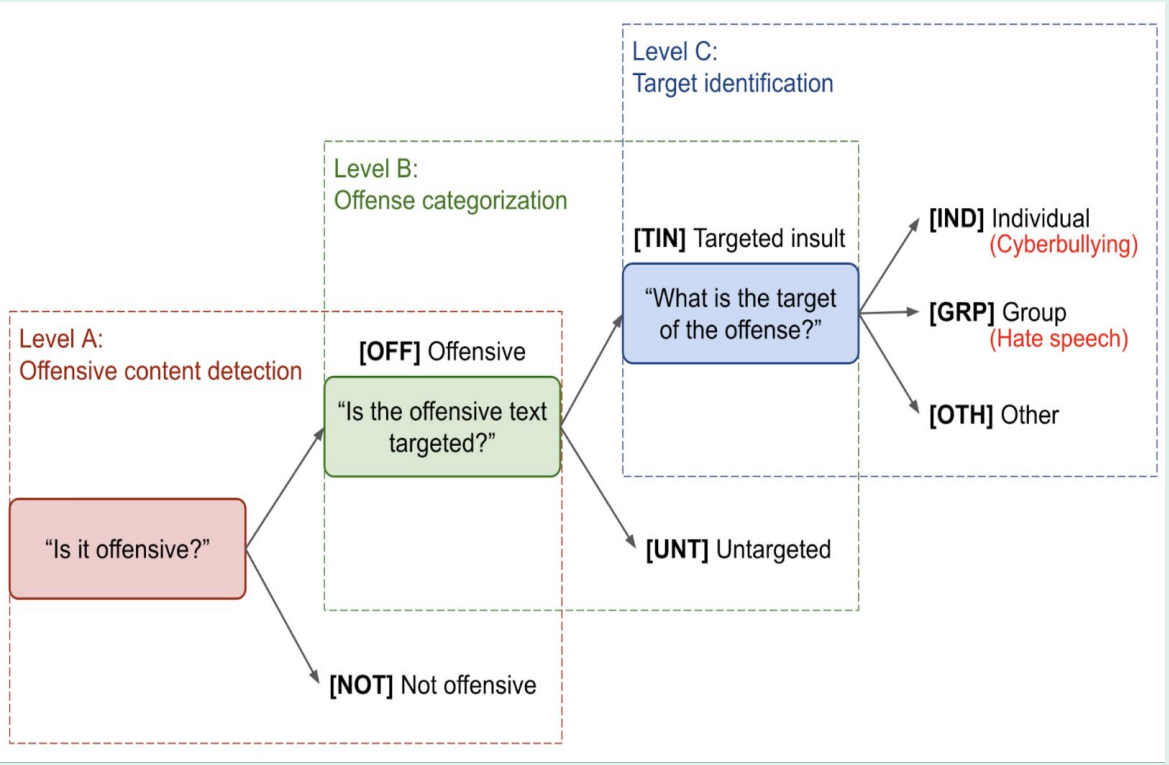
Transforming Online Presence Management

Our solution leverages the power of Sentiment Analysis AI to efficiently manage online presence. By integrating this innovative technology, businesses and individuals can swiftly identify and address hate speech and offensive language on social media platforms. This streamlines online reputation management, saving valuable time and resources. Acting as a mobile PR team, our solution aligns content with societal values, enhancing effective communication and audience engagement. With comprehensive analysis, businesses can safeguard their reputation, prevent mass media disruptions, and mitigate the detrimental effects of cancel culture. It empowers businesses to navigate the digital landscape with confidence, fostering a positive and inclusive online environment.

Pain points being solved

Identifying hate speech, Saving time and resources, Preventing reputational damage





Our Unique approach 💡

This innovative approach would provide businesses with a streamlined means to manage their online presence, saving considerable time and resources. Moreover, for individuals looking to create content, this AI-driven approach can act as a mobile PR team, assisting them in creating engaging and meaningful content while adhering to societal values.

Another significant advantage of this approach is that it would help break down discriminatory boundaries between people interacting online. By identifying and addressing hate speech and offensive language promptly, businesses can foster an inclusive and respectful environment for their audiences.

Application in Social Media Management

A. Assisting businesses in identifying acceptable content
   1. Determining social media guidelines and policies
   2. Establishing thresholds for hate speech detection

B. Preventing mass media disruption
   1. Alerting businesses to potential controversies or backlash
   2. Prompting proactive measures to address concerns

C. Combating cancel culture
   1. Helping businesses understand public sentiment
   2. Enabling timely responses and damage control strategies

Utilizing SocioSense enhances social media management, prevents disruptions, and supports businesses in effectively navigating online environments and shaping public opinion.

Identifying and Categorizing Speech

A. Utilizing sentiment analysis AI to replicate PR team's work
   1. Analyzing language patterns and emotions
   2. Identifying positive, negative, and neutral sentiments

B. Distinguishing hate speech from non-hate speech
   1. Recognizing discriminatory or offensive content
   2. Identifying harmful intentions or targeted attacks

C. Categorizing offensive hate speech
   1. Labeling content that incites violence or discrimination
   2. Identifying explicit or derogatory language

D. Categorizing non-hate speech
   1. Classifying content that promotes inclusivity and positivity
   2. Recognizing constructive criticism or dissenting opinions

Building the Path to Success

Building our concept

Once having our initial Idea and goal, the biggest hurdle was sourcing and cleaning the huge amounts of data required to train a model on the concept of hate speech which is quite contextual and distorted.

Finished

Stage 1

Development, testing, launch

Due to the complexities we faced at the training stage, development become a prolonged process of labelling and annotating data. the versatile nature of ate speech meant that we often relied on manually labelling data.
Our confusion matrix was still not optimal operating at about 60% accuracy once completed.

Finished

Stage 2

Future Developments

Although our prototype was a success and we were able to successfully demonstrate real world use cases and deployment capabilities, we still believe there is a vast room for improvement. We plan to focus on scalability and usability, allowing Sociosphere to not simply identify and evaluate but even educate users as well as aid in generating content ready for release from the ground up.

A work in progress

Stage 3

The team behind JustBePractical

Shuvam

Developer

A few lines of description and experience for each team member

Neel

Developer

A few lines of description and experience for each team member

The JustBePractical Solution: SocioSense

Unleash Online Success with AI Insights