

# Salud Mental

## PROYECTO 1 ETAPA 2

-Juan Carlos Marin  
-Kevin Cohen  
-Juan Felipe Castro



# Publicaciones de jóvenes en Reddit

## **SOBRE LA EMPRESA**

La empresa busca apoyar la detección de casos de posible suicidio mediante la información histórica de miles de usuarios de la plataforma Reddit, los cuales hacen publicaciones de diferentes tipos. Esto lo hacen con el fin de poder ayudar a esas personas que se encuentran en una situación crítica antes de que sea tarde, mediante un análisis automático que sea lo más rápido posible.



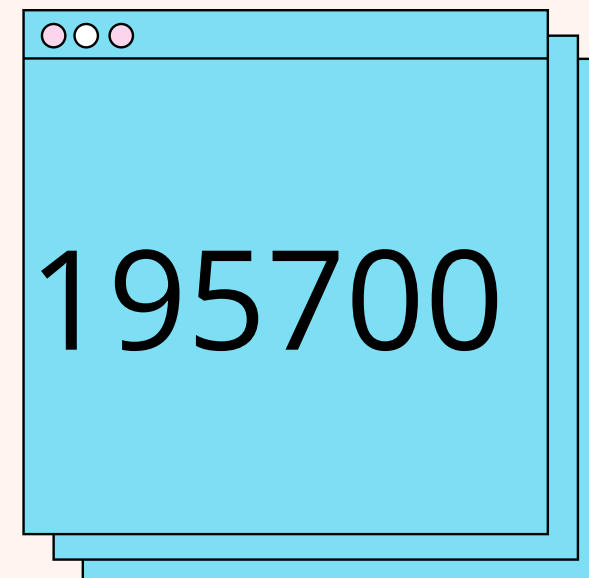
# ● Enfoque Analítico

## -Machine Learning para el modelo:



Lo que se busca es que mediante una base de datos que contiene las publicaciones de 195,700 usuarios y su destino final (suicidio o no) se construya un modelo de clasificación (mediante el procesamiento de lenguaje natural) que sea capaz de interpretar las publicaciones de futuros usuarios y determinar con cierto grado de precisión si esa persona tiene o no tendencias suicidas.

Total de publicaciones:



# Creación del PipeLine:

Para esta parte del proyecto utilizamos la base del modelo creado para la primer iteración, añadiéndole esta vez la construcción de un Pipeline capaz de consumir todas las funcionalidades a travez de un API que pueda ser usados por los usuarios finales:

```
In [21]: #Usamos la Libreria joblib  
filename = 'Proyecto1-Parte2.joblib'  
  
dump(pipeline, filename)
```

```
Out[21]: ['Proyecto1-Parte2.joblib']
```

```
In [28]: #Pasos que ha de seguir el pipeline  
transformer = ColumnTransformer(  
    [('vec', TfidfVectorizer(), 'text')], # column should be a string or int  
    remainder='passthrough'  
)  
  
pre = [('texto', Texto_Basico()),  
       ('token_y_mas', Texto_Definitivo()),  
       ('tfidf_vectorizer', transformer),  
       ('model', naive_bayes.MultinomialNB())]
```

```
In [29]: # Se ajusta el pipeline a los datos  
X = df_suicide.drop('class',axis=1)  
y = df_suicide['class']  
  
pipeline = Pipeline(pre)  
pipeline = pipeline.fit(X,y)
```

## Cambios en el codigo:

Retiramos dos  
algoritmos de  
clsificación

Algunas funciones  
de transformación



# Modelado y Evaluación

Algoritmos de clasificación utilizado:

## Naive Bayes



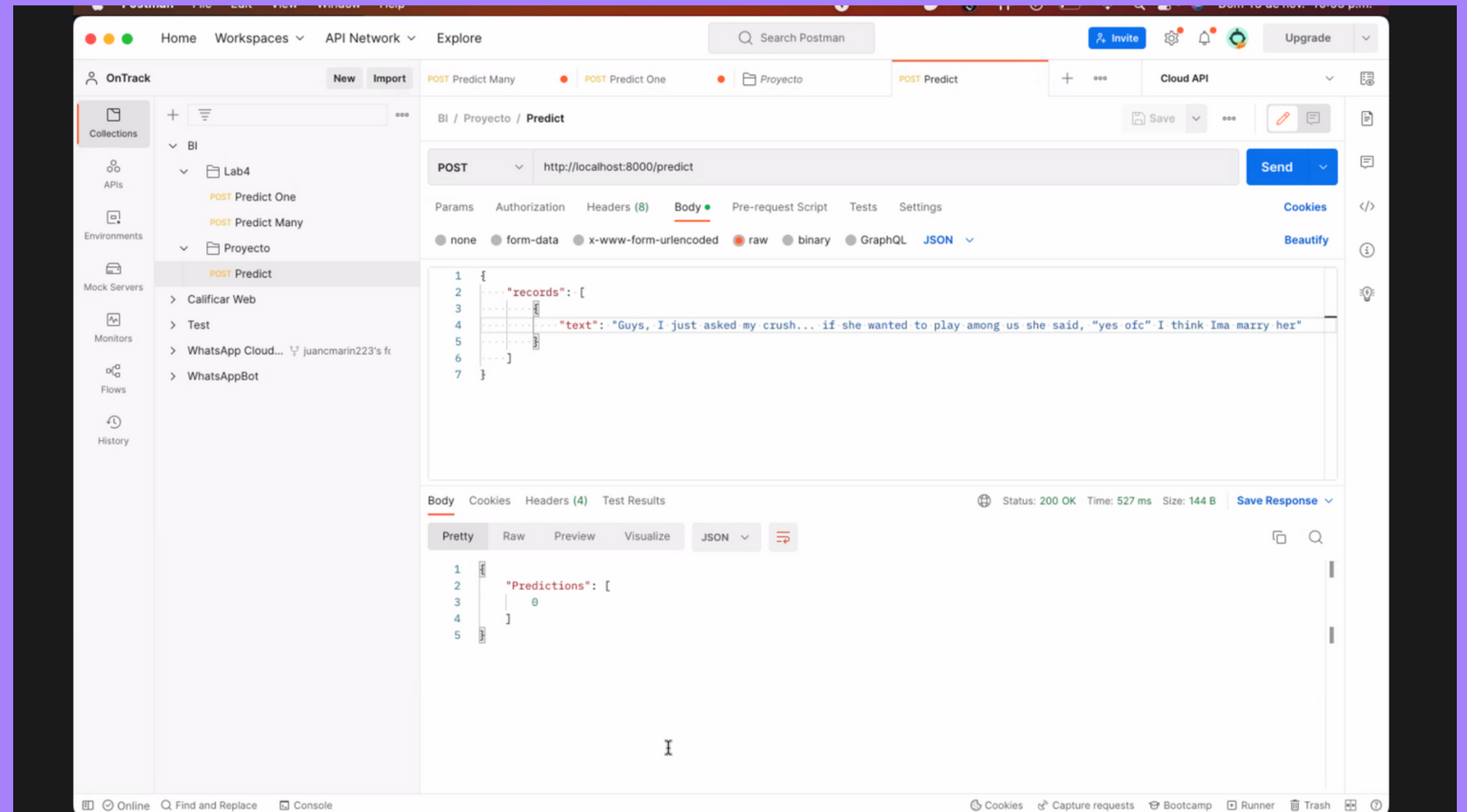
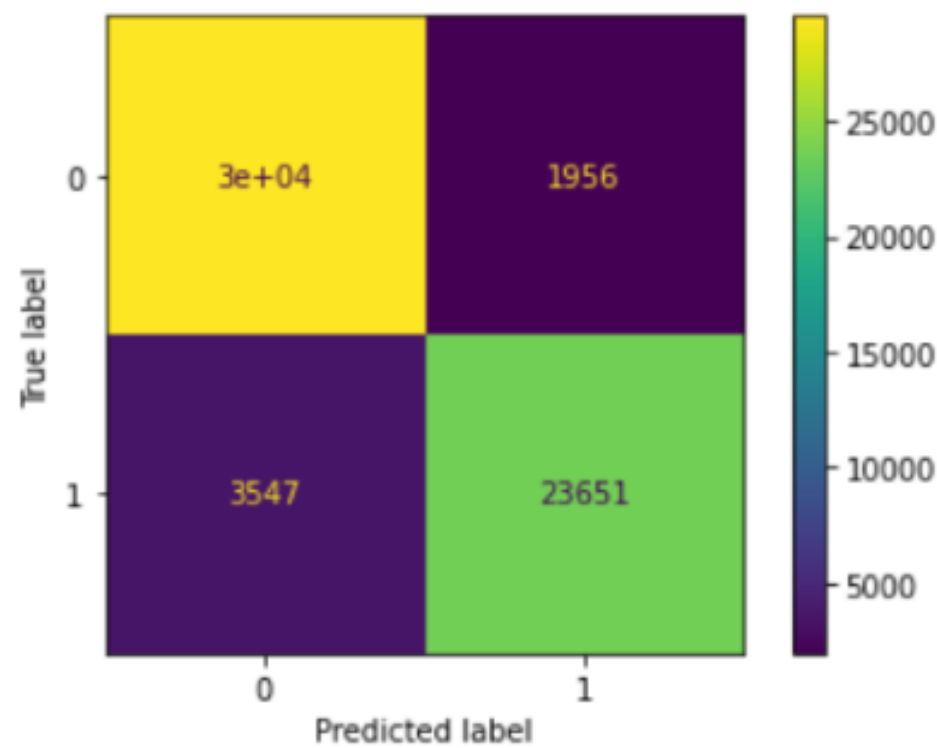
Este algoritmo aplica el teorema de Bayes y probabilidad condicional según las palabras tokenizadas y su influencia que tienen en la decisión de clasificación.

Solo dejamos implementado este algoritmo debido a los largos tiempos de ejecución que se presnetaron con los otros algoritmos.

# Analisis de resultados

## Naive Bayes

- Accuracy Score: 90.62
- Precision Score: 92.36
- Recall Score: 86.95
- f1 Score: 89.57



Para la prueba, introducimos un texto sin relación al suicido, y se obtuvo un muy buen resultado (0)

# Utilización de la aplicación

Potential Suicides Prediction from Reddit Posts

Reddit Post Info

Post Text:

"I kinda got behind schedule with learning for next week testweek I have 8 tests  
4 of them ive already studied for  
2 i studied good

Enviar

Possible No Suicidio

Potential Suicides Prediction from Reddit Posts

Reddit Post Info

Post Text:

i want to destroy myselffor once everything was starting to feel okay again but it all came tumbling down and i don't know how i used to cope before and the only reason i'm not tearing my skin to shreds and swallowing every pill i can find right now is because i'm not alone in the room but my walls are slowly falling down and it's only a matter of time before i snap and finally end it once and for all

Enviar

Possible Suicidio

# Posibles aplicaciones para el usuario

01

ESCUELAS: Los escuelas podrían utilizar esta información para entender que pasa por la cabeza de sus estudiantes y así ayudarlos mas fácilmente.

02

PSICOLOGOS: Estos modelos serian de gran ayuda para estos doctores, ayudando a identificar mas fácilmente posibles casos de suicidio, con tan solo algu.

03

ESCUELAS: Los colegios podrían beneficiarse de esta información compartiéndola con los padres de sus estudiantes, para que puedan monitorear sus redes sociales.

04

REDES SOCIALES: Estas plataformas podrian adaptar estos modelos asus algoritmos, para que dentro de la misma se detecten estas situaciones.



# Conclusiones

El modelo de esta iteración solo utiliza Naive Bayes

nuestro modelo esta en la capacidad de ayudar a las personas a detectar posibles casos de suicidio

El modelo necesita siempre estar acompañado de la opinion de un experto

El modelo no remplaza la labor de los especialistas..

