UNIVERSIDAD POLITÉCNICA SALESIANA

SEDE CUENCA

CARRERA DE COMPUTACIÓN

DESARROLLO DE UNA APLICACIÓN HÍBRIDA INTELIGENTE Y UN MODELO DE ANÁLISIS DE DATOS EN FASES PARA RESUMIR Y REPRESENTAR LOS COMENTARIOS DE CLIENTES SOBRE PRODUCTOS Y SERVICIOS DE BARES Y RESTAURANTES MEDIANTE DJANGO, IONIC, ANGULAR, DEEP LEARNING Y REDES SOCIALES

Trabajo de titulación previo a la obtención del título
de Ingeniero en Ciencias de la Computación e
Ingeniera en Ciencias de la Computación

AUTORES: JUAN CARLOS BARRERA BARRERA

KATHERINE MICHELLE BARRERA BARRERA

TUTOR: ING. REMIGIO ISMAEL HURTADO ORTIZ, PHD.

Cuenca - Ecuador

2022

CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN

Nosotros, Juan Carlos Barrera Barrera con documento de identificación N° 0106113301 y Katherine Michelle Barrera Barrera con documento de identificación N° 0106114309; manifestamos que:

Somos los autores y responsables del presente trabajo; y, autorizamos a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Cuenca, 26 de julio del 2022

Atentamente,

Juan Carlos Barrera Barrera	Katherine Michelle Barrera Barrera
0106113301	0106114309

CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA

Nosotros, Juan Carlos Barrera Barrera con documento de identificación N° 0106113301 y Katherine Michelle Barrera Barrera con documento de identificación N° 0106114309, expresamos nuestra voluntad y por medio del presente documento cedemos a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que somos autores del Proyecto Técnico: "Desarrollo de una aplicación híbrida inteligente y un modelo de análisis de datos en fases para resumir y representar los comentarios de clientes sobre productos y servicios de bares y restaurantes mediante Django, IONIC, Angular, Deep Learning y Redes Sociales", el cual ha sido desarrollado para optar por el título de: Ingeniero en Ciencias de la Computación e Ingeniera en Ciencias de la Computación, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribimos este documento en el momento que hacemos la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Cuenca, 26 de julio del 2022

Atentamente,

Juan Carlos Barrera Barrera	Katherine Michelle Barrera Barrera
0106113301	0106114309

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Remigio Ismael Hurtado Ortiz con documento de identificación N° 0104621388, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: DESARROLLO DE UNA APLICACIÓN HÍBRIDA INTELIGENTE Y UN MODELO DE ANÁLISIS DE DATOS EN FASES PARA RESUMIR Y REPRESENTAR LOS COMENTARIOS DE CLIENTES SOBRE PRODUCTOS Y SERVICIOS DE BARES Y RESTAURANTES MEDIANTE DJANGO, IONIC, ANGULAR, DEEP LEARNING Y REDES SOCIALES, realizado por Juan Carlos Barrera Barrera con documento de identificación N° 0106113301 y por Katherine Michelle Barrera Barrera con documento de identificación N° 0106114309, obteniendo como resultado final el trabajo de titulación bajo la opción Proyecto Técnico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Cuenca, 26 de julio del 2022

Atentamente,

Ing. Remigio Ismael Hurtado Ortiz, PhD.

0104621388

RESUMEN

Debido a la transformación digital aplicada en la mayoría de los negocios, para obtener una ventaja frente a la competencia, es necesaria la creación de herramientas inteligentes. Este proyecto está centrado en solventar la toma de decisiones del administrador de un bar o restaurante que ayuden al administrador a comprender su negocio y lograr alcanzar las metas propuestas.

La aplicación híbrida contiene el módulo de usuario con funciones para crear comentario, listar sus comentarios, y su perfil, de igual forma para el módulo de administrador se tiene funciones para el reporte con gráficos y explicaciones acerca de estado del restaurante o bar, según comentarios de los clientes.

El modelo de análisis fue separado en tres fases, la primera es limpieza del dataset junto con un análisis exploratorio, posterior a ello la segunda fase se enfoca en el entrenamiento y generación del modelo y finalmente la última fase contiene al método de predicción para la generación de un Resumen de texto ingresado. El dataset empleado para el entrenamiento del modelo fueron 200 comentarios que contienen 50% de comentarios positivos sobre el restaurante y 50% negativos.

Se integra este modelo dentro de la aplicación híbrida con la generación de archivos que contienen pesos, configuración y el tokenizador del modelo generado, esto permite la implementación de la Red Neuronal Transformer como un servicio. Cabe recalcar que la recolección de comentarios se obtiene de la aplicación hibrida y de la red social Facebook en la que se maneja la página oficial del restaurante, con esto se obtiene 25 publicaciones que se almacenan los comentarios de cada publicación categorizada por tipo de servicio y producto.

Finalmente, se aplicaron medidas de calidad para observar la precisión del modelo referente a la realización de un resumen. Se obtuvo un promedio de 0.36 entre Rouge 1,2 y L. Con base a esto se concluye que se obtiene un resumen abstractivo de un comentario, sin embargo, el modelo puede presentar valores atípicos cuando el texto de entrada es demasiado corto, por lo que se puede aplicar una solución en lo posterior como un trabajo futuro.

Palabras clave

Aplicaciones hibridas, Machine Learning, Resumen de texto, PLN, Transformers, T5, Redes sociales

ABSTRACT

Due to the digital transformation applied in most businesses, to gain an advantage over the competition, the creation of intelligent tools is necessary. This project is focused on solving the decision making of the manager of a bar or restaurant that help the manager to understand his business and achieve the proposed goals.

The hybrid application contains the user module with functions to create comments, list their comments, and their profile, in the same way for the administrator module it has functions for the report with graphics and details about the status of the restaurant or bar according to comments from the clients.

The analysis model was separated into three phases, the first is cleaning the dataset together with an exploratory analysis, after that the second phase focuses on the training and generation of the model and finally the last phase contains the prediction method for the generation of a Summary of entered text. The dataset used for model training was 200 comments containing 50% positive comments about the restaurant and 50% negative ones.

This model is integrated into the hybrid application with the generation of files containing weights, configuration and the tokenizer of the generated model, this allows the implementation of the Transformer Neural Network as a service. It should be noted that the collection of comments is obtained from the hybrid application and from the Facebook social network in which the official page of the restaurant is managed, with this 25 publications are obtained that store the comments of each publication categorized by type of service and product.

Finally, quality measures will be applied to observe the precision of the reference model to the realization of a summary. An average of 0.36 was obtained between Rouge 1,2 and L. Based on this, it is concluded that an abstract summary of a comment is obtained, however, the model may present outliers when the input text is too short, so a solution can be applied later as future work.

Key words

Hybrid Applications, Machine Learning, Text Summary, PLN, Transformers, T5, Social Networks,

Índice de Contenidos

RESUMEN 5
ABSTRACT 6
1. Introducción 15
2. Problema 16
2.1. Antecedentes 16
2.2. Importancia y Alcances 19
3. Objetivos Generales y Específicos 21
3.1. Objetivo General 21
3.2. Objetivos Específicos 21
4. Revisión de la literatura o Fundamentos Teóricos 22
4.1. Desarrollo de Software 22
4.1.1. Front-End 22
A. Angular 22
B. IONIC 22
4.1.2. Back-End 23
A. Django 23
B. Django REST Framework 23
C. Firebase 23
4.1.3. Servicios Web 23
4.2. Inteligencia Artificial 23
4.2.1. Machine Learning (Aprendizaje automático) y Reconocimiento de Patrones 24
Tipos de aprendizaje 25
1. Aprendizaje Supervisado 25
2. Aprendizaje No Supervisado 28
3. Aprendizaje Por Refuerzo 30
4. Aprendizaje Híbrido 31
5. Deep Learning 31
Redes Neuronales 32
1. Red neuronal recurrentes (RNN) 32
1.1. Redes de memoria a largo plazo (LSTM) 32
1.2. Unidad Recurrente Cerrada (GRU) 33
2. Redes neuronales convolucionales (CNN) 33

3. Redes Neuronales Transformers 33
4.2.2. IA basada en lógica matemática 33
4.2.3. Ingeniería de Búsqueda (Search) 34
4.2.4. Representación de Conocimiento 34
4.2.5. Planning 34
4.2.6. Heurísticas 34
4.2.7. Programación Genética (Genes) 34
4.3. Ciencia de datos 35
Metodología de análisis de datos CRISP-DM 35
Actividades de la Metodología CRISP-DM 36
1. Entendiendo el negocio 36
2. Recopilación y Comprensión de datos 36
Fuentes de Datos (Dataset) 36
2.1. IoT 36
2.2. Redes Sociales 36
3. Preparación de datos 37
Procesamiento del lenguaje natural 37
4. Elección de modelo 37
4.1. Resumen de texto 37
4.2. Análisis de datos 38
Análisis y Clasificación de sentimientos 38
5. Modelado 38
6. Despliegue o implementación 38
4.4. Calidad de servicio y satisfacción del cliente 39
4.5. Metodología de desarrollo Scrum 39
5. Marco Metodológico 42
5.1. Propuesta de Solución 42
5.2. Metodología 45
5.2.1. Modelo de Análisis aplicando metodología CRISP-DM 45
A. Entendiendo el negocio 45
B. Recopilación y Comprensión de datos 46
C. Preparación de datos 46
D. Elección de modelo 46
5.2.2. Metodología SCRUM 48

6. Resultados 54
6.1. Especificación de Requerimientos 54
6.1.1. Requerimientos Funcionales 54
6.1.2. Requerimientos No Funcionales 60
6.1.3. Diseño de Wireframes o Prototipos 61
6.2. Diseño de Alto Nivel 65
6.2.1. Vista de Desarrollo 65
6.2.2. Vista física 66
6.3. Diseño Detallado 67
6.3.1. Diagrama de Actividades 67
1. Inicio de Sesión 67
2. Registro de Usuarios 68
3. Creación de Comentario 69
4. Listar comentarios para Administrador 70
5. Reporte comentarios Administrador 71
6. Reporte comentarios extraídos de redes sociales 72
6.3.2. Diagrama de Clases 73
6.4. Desarrollo del Modelo de Análisis 74
6.4.1. Modelado 74
A. Fase 1. Preparación de Datos 74
B. Fase 2. Modelado 77
C. Predicción con nuevos samples 83
6.4.2. Implementación y despliegue como servicio 86
6.5. Codificación 86
6.5.1. Modelo Análisis 87
6.5.2. Aplicación Híbrida FrondEnd 88
6.5.3. Aplicación Hibrida BackEnd 90
6.5.4. Codificación de Web Services 93
6.6. Pruebas Unitarias 97
7. Cronograma 104
8. Presupuesto 107
9. Conclusiones 108
10. Recomendaciones 112
Referencias Bibliográficas 113

Anexos 118

Índice de Tablas

Tabla 1. Requerimiento para inicio de sesión 54
Tabla 2. Requerimiento para registrar usuarios 55
Tabla 3. Requerimiento para la creación de usuario/cliente 55
Tabla 4. Requerimiento para listar comentarios cliente/usuario 56
Tabla 5. Requerimiento para editar información básica del usuario 56
Tabla 6. Requerimiento para listar comentarios del administrador 57
Tabla 7. Requerimientos para generar reportes de comentarios 57
Tabla 8. Requerimiento para recolección de datos de redes sociales 58
Tabla 9. Requerimiento para la creación del modelo para resumen de comentarios 58
Tabla 10. Requerimiento para preparación de datos 59
Tabla 11. Requerimiento para creación del modelo de resumen de texto 59
Tabla 12. Requerimiento para predicción de nuevos valores 59
Tabla 13. Requerimientos no funcionales para la aplicación 60
Tabla 14. Tabla parámetros para entrenamiento de modelo 78
Tabla 15. Graficas de comportamiento a través de epochs y steps 80
Tabla 16. Parámetros para predicción de nuevos comentarios con modelo T5 84
Tabla 17. Prueba Unitaria para inicio de sesión 97
Tabla 18. Prueba Unitaria para Registro de usuario/cliente 98
Tabla 19. Prueba unitaria para creación de comentario de usuario 99
Tabla 20. Prueba unitaria para listar comentarios de usuario cliente 100
Tabla 21. Prueba unitaria para listar comentarios del administrador 101
Tabla 22. Prueba unitaria para generación de reportes de comentarios para el administrador 102
Tabla 23. Prueba unitaria de recolección de datos de redes sociales 103

Índice de Figuras

Figura 1. Estructura de la Aplicación 42
Figura 2. Proceso de Red Neuronal Transformers para resumen de texto 43
Figura 3. Modelo de adquisición de datos a través de Redes Sociales 44
Figura 4. Estructura del modelo Transformers T5 realizado por (Alammar, 2018) 47
Figura 5. Inicio de Sesión 62
Figura 6. Registro de Usuarios 62
Figura 7. Ingreso de Datos adicionales 62
Figura 8. Perfil de Usuario 63
Figura 9. Lista de Comentarios 63
Figura 10. Agregar nuevo comentario 63
Figura 11. Pagina frecuencia de palabras 64
Figura 12. Nube de palabras 64
Figura 13. Pagina para listar el resumen de comentarios 64
Figura 14. Arquitectura física de la aplicación hibrida 65
Figura 15. Diagrama de Despliegue de la aplicación hibrida 66
Figura 16. Diagrama de Actividades modulo Inicio de Sesión 67
Figura 17. Diagrama de Actividades modulo Registro 68
Figura 18. Diagrama de Actividades modulo creación de Comentario 69
Figura 19. Diagrama de Actividades modulo listado de comentarios 70
Figura 20. Diagrama de Actividades modulo Reportes de Comentarios 71
Figura 21. Diagrama de Actividades módulo de reportes de comentarios extraído desde redes sociales 72
Figura 22. Diagrama de Clase de la Aplicación enfocada a Comentarios. 73
Figura 23. Modelo en fases para construcción de Transformer enfocado en el resumen de texto 74
Figura 24. Palabras con más apariciones dentro del dataset 75
Figura 25. Palabras con más apariciones dentro del dataset una vez realizado un procesamiento de stopwords 75
Figura 26. Bigrama del Dataset sin aplicar stopwords 76
Figura 27. Bigrama del Dataset aplicando eliminación de stopwords 76
Figura 28. Grafica de perdida de entrenamiento (train los vs epochs) 80
Figura 29. Grafica de perdida de validación (val los vs epochs) 80
Figura 30. Grafica de perdida de entrenamiento (train los vs step) 80

Figura 31. Grafica de perdida de validación (val los vs steps) 80
Figura 32. Rouge-N del Modelo generado para resumen de texto 82
Figura 33. Resúmenes generados por nuestro modelo aplicando Transformer T5 83
Figura 34. Resúmenes generados por el Modelo T5 85
Figura 35. Parámetros servicio rest para consumo de modelo Transformer 86
Figura 36. Estructura organizacional de archivos del algoritmo 87
Figura 37. Requerimientos para desarrollo de FrontEnd con IONIC 6 88
Figura 38. Estructura del FronEnd de la aplicación implementada sobre IONIC 89
Figura 39. Configuración para Firebase 91
Figura 40. Configuración para Django y Facebook API 91
Figura 41. Estructura de Backend de la aplicación sobre Django 92

Índice de Anexos

A. Aplicación Móvil 118
A. Aplicación Móvil 117
Usuario 117
Administrador 120
B. Aplicación Web 122
Usuario 122
Administrador 127
C. Gráficas generadas desde Django para análisis estadístico 129
D. Codificación 131
Código de algoritmo para Resúmenes de texto 131
Código del FrontEnd para aplicación híbrida inteligente 131
Código del BackEnd para aplicación híbrida inteligente 131

1. Introducción

En la actualidad los negocios se presentan a grandes competencias tanto a nivel local como a nivel mundial, debido al avance tecnológico existente. Por lo que los clientes son cada vez más exigentes con el producto y su presentación. Esto se ha desarrollado con mayor rapidez después de la pandemia por COVID-19, debido a estas razones se consideró la transformación digital dentro de los negocios. En este caso específicamente hacia bares y restaurantes.

Debido a la facilidad de crear datos en esta época, se planteó el estudio de los comentarios existentes de un restaurante, tanto dentro de la aplicación desarrollada como de la red social de Facebook. Después de obtener los comentarios y clasificarlos por sentimientos posteriormente se obtiene un resumen por comentario mediante Redes Neuronales Transformers. Todo esto es presentado mediante un reporte con gráficas y explicaciones hacia el gerente o administrador del Restaurante o Bar para la toma de decisiones dentro del negocio. Esto será una herramienta que pueda proporcionar al administrador una idea del camino que debe tomar para generar mayor ingreso y un potencial crecimiento frente a su competencia.

El documento expone detalladamente el proceso que se realizó para la creación de la aplicación, desde una revisión de la literatura y los fundamentos teóricos adquiridos para el desarrollo, posterior un marco teórico que expone la propuesta y metodología aplicados, un apartado para los resultados obtenidos posterior al desarrollo finalmente se expone el cronograma actividades, el presupuesto, conclusiones y recomendaciones.

2. Problema

En esta sección se presenta los antecedentes que se toman en cuenta para la realización del proyecto, posterior a ello se expone la importancia y alcance para el problema planteado.

2.1. Antecedentes

En esta sección se presenta como actualmente la tecnología avanza de manera exponencial, esto conlleva a implementar mejoras o modificaciones sobre los servicios y productos que se ofertan hacia un público cada vez más exigente.

- La economía digital

  El avance tecnológico da acceso a nuevos mercados que aseguran una mayor demanda por parte de clientes, dando oportunidades de una mejor competitividad a pequeñas y medianas empresas. Chile es el primer país en Latinoamérica que cumple con los parámetros esenciales como; apertura de datos, confianza y emprendimiento digitales para la vinculación con la digitalización y el segundo lugar lo ocupa Colombia (Acosta, Yagual Velastegui, & Coronel Pérez, 2018).

- Impacto en empresas por COVID-19

  Después de una pandemia que marco un cambio radical como la forma de relacionarse en los periodos extendidos de confinamiento a creado una nueva realidad. Por lo que según el World Bank (Ungerer, Portugal, Molinuevo, & Natasha, 2020) el eCommerce tuvo un gran crecimiento porque reducía el riesgo de contagio por contacto y se ajustaba a todas las medidas de bioseguridad establecidas durante la pandemia. Por lo que una empresa debe evolucionar tanto en modelos de gestión como en estrategias de venta para ser sostenible. Como lo expone (Silvana Dakduk & Ricardo Dicarlo, 2020) se estimaba que para 2023 a nivel mundial el retail en eCommerce representaría el 21% de compras siendo el 73% compras por dispositivos móviles, y después de la pandemia se estima un crecimiento de hasta el 30% durante el año 2022 (Espinoza Vélez & Armijos Buitrón, 2022).

- Ventas en Comercio electrónico a nivel mundial

  Según señala (StatistaResearchDepartment, 2021) declara que las ventas de comercio electrónico en América Latina registran un valor de 70 000 millones de dólares en 2019 con una estimación de incremento del 66% para el 2023,

destacando países como Brasil, México, Colombia, Chile y Argentina. Se manifiesta según la Unión Internacional de Telecomunicaciones que, durante el 2020, el 49% de la población mundial son usuarios de Internet, 67% poseen un móvil y el 59% son parte de los usuarios de redes sociales. Además, se muestran datos expuestos por Internet World Stats y la Comisión Económica para América Latina y el Caribe (CEPAL) en el que Latinoamérica tuvo un aumento del 43,4% al 71,5% de usuarios con acceso a Internet, que supera el aumento mundial que es un promedio de 62% (Espinoza Vélez & Armijos Buitrón, 2022).

- Comercio electrónico en Ecuador

El país contaba durante el 2020 con 69% de personas con acceso a internet, el 80% de la población ecuatoriana contaba con un teléfono móvil y el 69% formaban parte de los usuarios en redes sociales, que incremento significativamente para el año 2021 en el que se presenta un aumento al 80.1% de personas con acceso a internet y 78.7% para los usuarios de redes sociales (Silvana Dakduk & Ricardo Dicarlo, 2020).

Por lo que es importante para las empresas ecuatorianas observar los cambios que ha tenido el consumidor para buscar alternativas de marketing digital que la posicionen como una competencia comercial. Según la (Silvana Dakduk & Ricardo Dicarlo, 2020) el 71% de las compras en línea son dirigidos a servicios como alojamientos y tiques, productos electrónicos y entretenimiento como música y juegos (Buenaño, Barriga, & Cadena, 2015). Por otro lado, las redes sociales que destacaron entre los ecuatorianos durante el periodo 2020 y 2021 según (Del-Alcázar, 2021) son Facebook con 13.3 millones, en segundo lugar, Instagram con 5.2, seguido de Spotify y Linkedin con 2.9. Otras cifras de carácter relevante son los eCommerce como Mercado Libre que tuvo un aumento de 1.7 millones de compradores nuevos y la entrega a domicilio tuvo un crecimiento de 1.1 millones (MercadoLibre, 2020). Según (Del-Alcázar, 2021) el 87% de compras se realizan vía web y el 85% a través de una app propia de la empresa o marca. El 49% prefiere comprar por la red social de Whatsapp, 44% por aplicaciones y 35% por sitios web. (Espinoza Vélez & Armijos Buitrón, 2022)

- Servicios de Comida

Según MINTUR (Ministerio de Turismo de Ecuador, 2019) existe 24.382 establecimientos registrados en Ecuador de los cuales el 85% son alojamientos y servicios de comida hasta el 2019, con un crecimiento de 2,9% entre los años 2011 y 2019. Según el sistema SAIKU del SRI las ventas fueron un total de 1.919,17 millones en 2019 con un crecimiento del 6,9% entre 2011 y 2019 (Espinoza Vélez & Armijos Buitrón, 2022). De acuerdo con la base de datos de la Zona 6 del Ministerio de Turismo existían un total de 1.400 establecimientos de alimentos y

bebidas que están registrados y siguen activos en Cuenca, se compone de 32 discotecas, 155 bares, 850 restaurantes y el restante lo conforman cafeterías y fuentes de soda (Ministerio de Turismo, Ecuador, 2016).

Pero al igual que varios negocios estos también fueron afectados de manera significativa durante la pandemia, inclusive con la creación de 'dark kitchens', también conocidas como cocinas fantasmas o virtuales, son un nuevo modelo de restaurante en tiempos en que las personas deben limitar sus salidas a la calle por la pandemia de COVID-19 (Coba, 2020). Lo que ha creado algunos restaurantes enfocados en un aplicación móvil o redes sociales para tener la interacción con su cliente para finalizar una compra y recibir el pedido en el domicilio especificado por el cliente. (Reinoso Naranjo, 2021)

Resultados expuestos de una encuesta enfocada en los restaurantes existentes en Quito, los clientes se dividen en 4 grupos, el primer grupo que son el 13,7% con preferencia por salir a comer en un restaurante, 16,9% prefiere comprar comida para llevar, el 25,5% opta por la comida realizada en la casa y el 43,9% son personas que realizar un pedido a domicilio para obtener sus alimentos durante la pandemia (Reinoso Naranjo, 2021). Debido a esta situación este tipo de negocios han evolucionado y necesitan nuevas estrategias que mejoren la atención al cliente y la gestión para crear un negocio rentable y sostenible.

- Big Data

Conforme el crecimiento de dispositivos digitales usados en la actualidad como PC, tables, teléfonos inteligentes, sensores, entre otros, marcan un gran crecimiento en el volumen de datos que se generan día a día. (Penkler, 2018). Por lo que en los últimos años los negocios han cambiado debido a que encontraron el valor en los datos, marcado por las compras en línea de servicios o productos (Singha, Bilal, & Wornchanok, 2019). En la actualidad la mayor parte de actividades cotidianas realizadas por las personas conllevan algún tipo de rastro digital que puede contener información valiosa para conocer las necesidades que se deben solventar en el mercado como la resolución de estos problemas, en consecuencia, los datos tienen un alto valor comercial y social para las empresas de hoy en día (Jeffrey, Luca, Marco,, Emanuela, & Riccardo, 2018). Según IDC el pronóstico para el crecimiento de Big-Data es mayor al 50% entre 2015 a 2019. Con aumento de 122 mil millones a 187 mil millones (Gartner, 2016). Al igual que la encuesta realizada por Well Street Journal junto a Oracle expone que el 86% de los ejecutivos reconoce el gran beneficio que proporciona el Big-Data y es una de las principales prioridades al momento de tomar decisiones comerciales estratégicas. (Oracle, 2015)

- Análisis de Datos

El análisis de los datos es la actividad de transforma Big-Data en información de valor, mediante la aplicación de distintos algoritmos que depende lo que se desea analizar. La utilidad para las empresas está en predecir el comportamiento humano para la mayor probabilidad de éxito empresarial. (Sol & Narasimha, 2021). El análisis de Big Data es el descubrimiento de conocimientos, perspectivas e inteligencia a partir de estos datos. La visualización, informe de conocimiento y perspectivas descubiertas para ayudar con la toma de decisiones. (Z. Sun & Strang, 2018).

Los componentes principales son el análisis descriptivo que consiste en descubrir y explicar el problema de como sucedió, cuando y que está sucediendo, el siguiente componente es el análisis predictivo enfocado a pronosticar o predecir tendencias abordando el problema de que es probable que suceda y por qué sucederá. Y finalmente el análisis prescriptivo que aborda el problema de que se debe hacer y que es lo que debe suceder con el mejor resultado con un bajo índice de incertidumbre. (Z. Sun & Strang, 2018) (Sharda, Delen, & Turba). Esto vuelve al análisis de datos una de las herramientas más cotizadas por empresas porque ayuda a la toma de decisiones respecto a servicios o productos ofertados como la expansión comercial que debe tomar la empresa y encontrar tendencias con el menor margen de error posible. (Ruchi, Yashaswi, & Dr. Jongwook, 2017)

2.2. Importancia y Alcances

Dentro de la siguiente sección se profundiza cual es el grupo objetivo al cual está enfocado este proyecto, haciendo uso de una justificación para este motivo.

Justificación

Una solución para la evolución que vivieron los negocios ecuatorianos postpandemia, y con el auge de la transformación digital es indispensable la creación de herramientas inteligentes enfocadas en actividades que pueden ser automatizadas mediante programas o sistemas para mejorar y crecer en el mercado.

La herramienta inteligente que propone el presente trabajo está centrada en la opinión que los clientes pueden tener de un restaurante o bar al que acudieron, esto con la finalidad de mejorar la experiencia del servicio y la calidad producto ofertado por estos establecimientos. Además de la mejora enfocada a la atención del cliente, esta implementación aumenta las ganancias por ventas o servicios del restaurante y posición estratégica sobre sus competidores.

La automatización para interpretar y transformar la información expuesta por el usuario, en este caso su comentario acerca del servicio o producto es de suma importancia para obtener información que sea de alto valor para el administrador del establecimiento que puede en base a los datos obtenidos, analizados y estudiados optar por unas decisiones que presentan un porcentaje elevado de éxito. Al igual que la reducción de perdidas por factores externos o internos que generalmente son poco visibles sin el estudio previo de los datos manejados por el establecimiento.

Esto incrementa la posibilidad de asegurar una compra o la satisfacción con un servicio que oferta el restaurante, por lo que el marketing digital tiene una gran importancia al combinarse con las capacidades de las redes sociales de obtener información personal o comentarios acerca de articulo o servicio vendido. La interacción con el usuario y potencial cliente es vital para realizar una autoevaluación de la trayectoria que debe tomar la empresa o negocio a fin de solventar sus necesidades existentes en el mercado, mejoras y toma de decisiones que pueden ser aplicadas para tener preferencia sobre las demás empresas.

La herramienta inteligente expuesta en este trabajo consiste en una aplicación multiplataforma, el usuario podrá acceder a los servicios mediante el navegador web, o desde su teléfono inteligente. Contiene un modelo de análisis de datos con el fin de resumir y representar comentarios de los clientes de un bar o restaurante sobre el producto que consumió en el establecimiento como el servicio, conformidad con la ambientación y la atención que recibió durante su estadía en el establecimiento. Esto se realiza mediante tecnologías estándar utilizadas en la actualidad como el entorno de desarrollo web Django, IONIC, Angular y métodos de Deep Learning para el análisis de datos.

3. Objetivos Generales y Específicos

En la siguiente sección se presenta los objetivo general y objetivos específicos que se desea alcanzar con el presente proyecto.

3.1. Objetivo General

Desarrollar una aplicación híbrida inteligente y un modelo de análisis de datos en fases para resumir y representar los comentarios de clientes sobre productos y servicios de bares y restaurantes mediante Django, IONIC, Angular, Deep Learning y Redes Sociales.

3.2. Objetivos Específicos

- **OE1.** Estudiar y conocer los fundamentos de Deep Learning, procesamiento de lenguaje natural, IONIC, Angular, Django y herramientas de extracción de datos.

- **OE2.** Diseñar y desarrollar el modelo de análisis de datos en fases para resumir y representar comentarios.

- **OE3.** Diseñar y desarrollar la aplicación híbrida y el módulo de adquisición de datos de redes sociales.

- **OE4.** Integrar el modelo de datos como servicio dentro de la aplicación hibrida.

- **OE5.** Diseñar un plan de experimentación del modelo de análisis y de la aplicación híbrida

4. Revisión de la literatura o Fundamentos Teóricos

En esta sección, se abordará los conceptos de las tecnologías utilizadas para el planteamiento de este proyecto. En la primera sección se muestra el desarrollo de software, en la segunda sección se expone sobre la inteligencia artificial, la tercera sección trata sobre la ciencia de datos, como cuarta sección se tiene la calidad de servicio y satisfacción del cliente, finalmente la quinta sección muestra la metodología de desarrollo Scrum aplicada al proyecto.

4.1.Desarrollo de Software

En este apartado se explica las tecnologías que se manejan dentro de la aplicación, en la primera sección tenemos las tecnologías para el Front-End, en la segunda sección tecnologías para Back-End y finalmente una breve descripción de Servicio Web.

4.1.1. Front-End

A. Angular

Angular es uno de los marcos más populares de JavaScript y desarrollo web, este framework tiene un enfoque en el desarrollo de aplicaciones web Front-End, su mantenimiento y desarrollo es realizado por el equipo de Angular en Google. Su lenguaje de programación es TypeScript, un superconjunto de JavaScript, que hace uso de componente enfocado en módulos, dando soporte a programación orienta a objetos, carga dinámica, conjunto personalizado de componentes de interfaz de usuario, generadores estilo Python, lambdas e iteradores, entre otros (Uzayr, Cloud, & Ambler, 2019).

B. IONIC

IONIC Framework ofrece un conjunto de varias funciones predefinidas para trabajar en la interfaz de usuario de una aplicación, sin preocuparse por las diversas superficies de trabajo de la aplicación móvil. Una aplicación IONIC puede construirse a través de las dos tecnologías principales: Angular y Apache Cordova, permitiendo crear aplicaciones híbridas, que permite desarrollar una interfaz de aplicación, como si fueran páginas web, es decir, usando HTML, CSS, JavaScript. Estas aplicaciones se ejecutan dentro de WebView de la aplicación nativa (Chaudhary, 2018).

4.1.2. Back-End

A. Django

Es un entorno de desarrollo web, la interfaz de programación de aplicaciones (API) de Django es un mecanismo completo para la administración de todas las partes incluyendo base de datos y plantillas html. Django es un mecanismo que ofrece todas las herramientas para el montaje de aplicaciones web Full-Stack (Ashley, 2020).

B. Django REST Framework

Es un paquete que nos permite crear API WEB, que se integra a la perfección con Django que permite completar las funciones necesarias de un API REST, entre la variedad de componentes se tienen las vistas REST basadas en clases, conjuntos de vistas y serializadores (Gagliardi, 2021).

C. Firebase

Firebase se considera un backend como servicio, que ahora forma parte de Google Cloud Platform mientras sigue funcionando como una entidad independiente. Ofrece diferentes servicios como alojamiento, bases de datos en tiempo real y funciones en la nube (Hajian, 2019).

4.1.3. Servicios Web

También conocido como REST, representa un modelo de arquitectura de software que permite definir un conjunto de servicios o aplicaciones individuales simples o complejas a través de la web disponibles para todos los usuarios (González & Manuel, 2014).

4.2. Inteligencia Artificial

La Inteligencia Artificial (IA) es una de las innovaciones tecnológicas que sucedieron, para reemplazar el trabajo manual que realiza el ser humano en varios campos. La inteligencia artificial es una rama de la ciencia y la tecnología que crea máquinas inteligentes y programas informáticos para realizar diversas tareas que requieren inteligencia humana. Es un sistema que imita varias funciones que un ser humano

puede hacer. AI utiliza datos externos como big data para lograr un excelente rendimiento para las tareas dadas (Lavanya, 2021).

La Inteligencia Artificial, específicamente la Minería de Datos, Machine Learning (aprendizaje automático) y Deep Learning (aprendizaje profundo) son herramientas para ayudar a anticipar problemas operativos y tomar medidas prescriptivas para evitarlos.

4.2.1. Machine Learning (Aprendizaje automático) y Reconocimiento de Patrones

Es un conjunto de métodos que detectan patrones en los datos automáticamente y posterior a ello realizar una predicción de resultados futuros o distintos tipos de toma de decisiones bajo incertidumbre.

Aplicaciones de Machine Learning

- **Marketing:** Mejora la toma de decisiones.
- **Marketing Personalizado:** Encontrar nuevos mercados, realizar análisis de comportamientos y predicción ventas a futuro.
- **Detección de Fraude**
- **Reconocimiento de imágenes:** Reconocimiento de Manuscritos y Digitalización de textos.
- **Reconocimiento de rostros**
- **Bolsa de valores:** Predecir tendencias
- **Comercio Financiero:** Análisis financiero.
- **Atención al cliente:** Aplicar mejoras mediante Procesamiento de Lenguaje Natural (NLP) y chatbots

- **Predicción de enfermedades**
- **Ecomerce**
- **Redes basadas en Intención**
- **Evaluación de riesgos**
- **Cybersecurity:** Detección de Malware, Spam

Tipos de aprendizaje

1. **Aprendizaje Supervisado**

El modelo de aprendizaje supervisado pretende distinguir algunas partes de los datos para la clasificación de patrones con datos etiquetados (Kim, Aminanto, & Tanuwidjaja, 2018).

- **Clasificación**

En este tipo de aprendizaje, la variable desconocida es discreta, con esto llegando a generar grupos predefinidos. A través de este aprendizaje se puede llegar a reconocer imágenes, diagnóstico de patologías de exámenes médicos, identificación de rostros, entre otros.

- **Regresión**

Según (Ortiz, 2022), regresión "es el cálculo de relaciones matemáticas entre una variable continua y una o más variables. Una función de regresión permite estimar el valor de una variable dado el valor de otra, para los valores que no se han obtenido antes".

Tipos de Algoritmos supervisados

- **Naive Bayes**

Naive Bayes Classifier es un modelo de aprendizaje automático que se basa

en los supuestos de probabilidad condicional (Le, 2016). Naive Bayes se encuentra dentro de los clasificadores probabilísticos simples basados en la aplicación del teorema de Bayes con suposiciones fuertes de independencia entre las características (Vasudha Goswami, 2020).

Naive Bayes Classifier puede predecir si un nuevo mensaje de texto se categoriza como positivo, negativo o neutral. Se utiliza para predecir la probabilidad de que una palabra dada pertenezca a una clase en particular (Naing, Thwe, Mon, & Naw, 2018).

Primero, el clasificador de Naive Bayes calcula la probabilidad previa. En segundo lugar, calcula la probabilidad condicional de cada atributo de palabra. En tercer lugar, calcula la probabilidad posterior. Finalmente, Naïve Bayes Classifier determina la clase (Naing, Thwe, Mon, & Naw, 2018).

- **Máquinas de soporte vectorial (SVM)**

Support Vector Machine es un modelo de aprendizaje supervisado. Analiza los datos utilizados para la clasificación y el análisis de regresión que encuentra el hiperplano que diferencia las dos clases a predecir. SVM puede realizar problemas de clasificación tanto lineales como no lineales (Vasudha Goswami, 2020).

SVM encuentra un hiperplano lineal (límite de decisión) que separa las características de tal manera que se maximiza el margen. Y luego, el clasificador SVM determina el (positivo, negativo o neutral). Los nuevos datos de prueba se mapean en el mismo espacio y se predice que pertenecen a una categoría según el lado de la brecha en el que se encuentran (Naing, Thwe, Mon, & Naw, 2018).

- **K-Nearest Neighbor (KNN)**

Para llegar a entender de mejor manera este algoritmo se hace una inferencia de la vida real, en donde las personas tienden a tomar decisiones o emprender acciones basándose en los consejos de las personas que los rodean, especialmente de aquellos que están muy cerca de nosotros (Zhou, 2020). El método k-vecino más cercano (KNN), toma K y nos indica que esta decisión estará basada en k vecinos, entonces este algoritmo se basa en la idea de que los patrones más cercanos a un patrón de destino x, brindan información de etiqueta útil para este nuevo patrón, en donde se busca etiquetarlo (Kramer, 2013).

- **Arboles de decisión (Random Forest)**

Según (Le, 2016), "Un árbol de decisión es una herramienta de apoyo a la decisión que utiliza un gráfico o modelo de árbol de decisiones y sus posibles consecuencias, incluidos los resultados de eventos fortuitos, los costos de los recursos y la utilidad".

Random se refiere al muestreo aleatorio de datos del conjunto de original, y Forest se refiere a la construcción de árboles de decisión múltiples, uno para cada muestra aleatoria de datos. Por lo que, Random Forest crea múltiples árboles de decisión, donde cada árbol de decisión funciona en una muestra de los datos (Ayyadevara, 2018).

- **Regresión de mínimos cuadrados ordinarios (OLS)**

Es una técnica utilizada para explorar si una o múltiples variables (la variable independiente X) pueden predecir o explicar la variación en otra variable (la variable dependiente Y). La regresión OLS pertenece a una familia de técnicas denominadas modelos lineales generalizados, por lo que las variables que se examinan deben medirse a nivel de razón o intervalo y tener una relación lineal (Wooditch, Johnson, Solymosi, Ariza, & Langton, 2021). Para el entendimiento de OLS se toma una gráfica lineal y luego, para cada uno de los puntos de datos dentro del plano, se mide la distancia vertical entre el punto y la línea, para posteriormente sumarlos; la recta ajustada sería aquella en la que esta suma de distancias sea la menor posible (Le, 2016).

- **Regresión logística**

La regresión logística se utiliza para predecir la probabilidad de cada clase objetivo. En caso de clasificación, devuelve la probabilidad asociada a cada clase para cada registro (Singh, 2018).

La regresión logística utiliza la regresión lineal para capturar la relación entre las variables de entrada y salida, además se usa un elemento adicional como una función no lineal para convertir la salida de forma continua en probabilidad (Singh, 2018).

Según (Le, 2016), " La regresión logística mide la relación entre la variable dependiente categórica y una o más variables independientes mediante la estimación de probabilidades utilizando una función logística, que es la distribución logística acumulativa"

2. Aprendizaje No Supervisado

Dentro de este aprendizaje se hace uso de datos no etiquetados, a partir de los datos de entrenamiento limitados, es esencial aprender cada capa inferior en un enfoque capa por capa sin depender de todas las capas superiores. Estos modelos calculan la probabilidad conjunta, dada la entrada y eligen la etiqueta de clase con la probabilidad más alta. Los modelos generativos también pretenden aprender distribuciones estadísticas del conjunto de datos (Kim, Aminanto, & Tanuwidjaja, 2018).

- **Métodos de agrupamiento**

Estos se pueden ver como la detección automática de grupos de ejemplos que tienen características similares, que pueden indicar posiblemente el hecho de que un miembro del grupo pertenece a una clase bien definida.

- **Métodos de asociación**

Estos son un problema muy relevante para los comerciantes en línea, y consisten en detectar grupos de elementos que se observan con frecuencia en conjunto. Ejemplo: sugerir compras adicionales a un usuario, según el contenido de su carrito de compras.

Algoritmos de aprendizaje no supervisado

- **Clustering**

K-Means: Tiene la ventaja de que es bastante rápido, ya que todo lo que realmente estamos haciendo es calcular las distancias entre los puntos y los centros de los grupos (Seif, 2018). El clustering de K-Means segrega 'n' observaciones en 'k' conglomerados o grupos, donde cada objeto o punto de datos pertenece al conglomerado con la media más cercana. Para esto, elige los centroides iniciales 'k' arbitrariamente y los grupos se forman en función de la proximidad de los objetos de datos con estos centroides (Arora, Virmani, Jindal, & Sharma, 2017).

Mean Shift: Es un algoritmo basado en ventanas deslizantes (sliding-window) que intenta encontrar áreas densas dentro de los puntos de datos

(Seif, 2018). Las ventanas deslizantes forman una ventana sobre una parte de los datos, y esta ventana puede deslizarse sobre los datos para capturar diferentes partes de ellos (Moore, 2019). Este algoritmo se encuentra basado en el centroide, entonces su objetivo es ubicar los puntos centrales de cada grupo/clase, esto se logra actualizando los candidatos para que los puntos centrales sean la media de los puntos dentro de la ventana deslizante (Seif, 2018). Estas ventanas luego se filtran en una etapa de pos-procesamiento para eliminar duplicados, formando el conjunto final de puntos centrales y sus grupos correspondientes. A diferencia del agrupamiento de K-means, no es necesario seleccionar el número de agrupamientos, ya que el desplazamiento de la media lo descubre automáticamente (Seif, 2018).

K-Medoids: Es un método de partición de la forma clásica de agrupamiento en el que un conjunto de datos dado n. Es mucho más fuerte al ruido y su valor extremo se compara con k-means porque a pesar de tomar algunas de las distancias euclidianas al cuadrado, prefiere la suma reducida de las diferencias por pares (Sharma, Kumar, & Shrivastava, 2020).

- **Reducción de Dimensionalidad**

Análisis de componentes principales (PCA): Es un método matemático para evaluar los componentes principales de un conjunto de datos. Los componentes principales son un conjunto de vectores en un espacio de alta dimensión que capturan la varianza o la variabilidad del espacio de características (Bisong, 2019), PCA es una técnica estadística para transformar un conjunto de datos original a un modelo de menor cantidad de variables, mediante la composición de las características más importantes que capturan la máxima información sobre el conjunto de datos.

Selección por características (FS): La selección de funciones o características (FS) es importante en las tareas de aprendizaje automático porque puede mejorar significativamente el rendimiento al eliminar funciones redundantes e irrelevantes y, al mismo tiempo, acelerar la tarea de aprendizaje (Silva & Leong, 2015). Este método se basa en información mutua, que es una medida de relevancia y redundancia entre las otras funciones (Verma & Salour, 2020). FS ayuda principalmente a aliviar la dimensionalidad y acelerar la tarea de aprendizaje. En la práctica, también ayuda a optimizar las metodologías de recopilación de datos (Silva & Leong, 2015).

Análisis discriminante lineal (LDA): Describe las diferencias entre los grupos en términos de funciones discriminantes (DF), también llamadas

variables canónicas (CV), definidas como combinaciones lineales de las variables originales (Trendafilov & Gallo, 2021).

LDA intenta reducir las dimensiones del conjunto de características mientras retiene la información que discrimina en las clases de salida. LDA intenta encontrar un límite de decisión alrededor de cada grupo de una clase. A continuación, proyecta los puntos de datos a nuevas dimensiones de manera que los grupos estén lo más separados posible entre sí y los elementos individuales dentro de un grupo estén lo más cerca posible del centroide del grupo.

Las nuevas dimensiones se clasifican en función de su capacidad para maximizar la distancia entre los grupos y minimizar la distancia entre los puntos de datos dentro de un grupo y sus centroides. Estas nuevas dimensiones forman los discriminantes lineales del conjunto de características.

3. Aprendizaje Por Refuerzo

Según (Ortiz, 2022), dentro de este tipo de aprendizaje por “cada paso incorrecto el sistema es castigado y por cada paso correcto el sistema es recompensado”.

Algoritmos de aprendizaje por refuerzo

- **Aprendizaje de diferencia temporal**

Usamos los recuerdos/experiencias del pasado para hacer suposiciones informadas de lo que probablemente podría ser el futuro. El aprendizaje TD (por diferencia temporal) logra un buen equilibrio entre los enfoques basados en la programación dinámica que requieren "sin experiencia" y los enfoques basados en Monte Carlo que utilizan "solo experiencia" (Sewak, 2019).

El algoritmo de diferencia temporal proporciona un mecanismo en línea para el problema de estimación. la diferencia temporal podría ser adaptativa para usarse en un enfoque que sea similar a la programación dinámica o la simulación de Monte Carlo o cualquier cosa intermedia (Sewak, 2019).

- **Sarsa**

El acrónimo SARSA significa State Action Reward State Action. Utiliza el mismo principio para las actualizaciones de la función de valor y lo aplica a

las actualizaciones de la función de acción-valor. SARSA trabaja en el lado del "control" del problema. SARSA es un enfoque dentro de la política y usa la misma política para el comportamiento y para la estimación y requiere una selección cuidadosa de los valores de inicialización para evitar posibles inconvenientes (Sewak, 2019).

- **Q-Learning**

Q-Learning también utiliza el aprendizaje de diferencias temporales (TD Learning) para el lado de la estimación del problema. Además, al igual que SARSA, Q-Learning proporciona una solución para la parte de "control" del problema e intenta estimar el valor de acción/Función Q para tomar la mejor acción posible. Q-Learning utiliza las tuplas (estado-acción-recompensa-estado) como experiencia para estimar la Función Q (Sewak, 2019).

4. Aprendizaje Híbrido

La arquitectura profunda híbrida combina arquitecturas generativas y discriminativas. La estructura híbrida tiene como objetivo distinguir los datos, así como el enfoque discriminativo. Sin embargo, en el paso inicial, ha ayudado de manera significativa con los resultados de las arquitecturas generativas. Un ejemplo de arquitectura híbrida es Deep Neural Network (DNN) (Kim, Aminanto, & Tanuwidjaja, 2018).

5. Deep Learning

El aprendizaje profundo proviene originalmente de los avances del algoritmo el vecino más cercano (NN sus siglas en inglés), que aplica varios métodos para superar las limitaciones de una sola capa oculta en NN.

Estos métodos emplean capas ocultas consecutivas que están en cascada jerárquicamente, por lo tanto, el entrenamiento de redes neuronales con más de cinco capas no lineales se denomina aprendizaje profundo (Deep Learning). Haciendo que dentro de este apartado exista una variedad de modelos, clasificados en tres subgrupos, generativo (no supervisado), discriminativo(supervisado) e híbrido.

La clasificación se basa en la intención de las arquitecturas y técnicas, por ejemplo, síntesis/generación o reconocimiento/clasificación (Kim, Aminanto, & Tanuwidjaja, 2018).

Redes Neuronales

Según, (Salas, 2018) "Una red neuronal está basada en el funcionamiento en las redes neuronales biológicas que están presentes en el cerebro de una persona, la base principal son las neuronas y sus conexiones para poder simular el comportamiento humano. En el caso de una red neuronal artificial se compone primordialmente por una capa de entrada, una capa de salida y N capas ocultas. Cada una de las neuronas o también llamados nodos tienen un valor umbral para su activación en la fase de entrenamiento de la red neuronal".

Tipos de Redes Neuronales

1. **Red neuronal recurrentes (RNN)**

   Las redes neuronales recurrentes (RNN) son una extensión de las redes neuronales con enlaces cíclicos para procesar información secuencial. Estos enlaces cíclicos colocados entre las neuronas de capa superior e inferior permiten a RNN propagar datos de eventos anteriores a los actuales. Esta propiedad hace que RNN tenga una memoria de eventos de series temporales (Kim, Aminanto, & Tanuwidjaja, 2018).

   Según, (Zaremba, 2015) "Una red neuronal recurrente no consta con una distribución definida, para permitir las conexiones entre neuronas, incluso de modo que sea posible crear una temporalidad que permita que la red tenga memoria ya se a corto o largo plazo, siendo muy potentes para todo lo que tiene que ver con el procesamiento de sentimientos".

   Según, (Kurniasari & Setyanto, 2020) "Una aplicación práctica es la clasificación de sentimientos para la evaluación de reseñas y/o comentarios de algún producto o servicio y conocer si el comentario es positivo, negativo o neutral. Basado en los resultados anteriores se mide la satisfacción del cliente en cuanto al servicio proporcionado".

   En las redes neuronales recurrentes se presentan las siguientes clasificaciones:

   1.1. Redes de memoria a largo plazo (LSTM)

   Esencialmente, los LSTM funcionan típicamente como RNN, pero son capaces de aprender dependencias a largo plazo. Alivian el problema de la desaparición o explosión del gradiente y brindan una precisión mucho mayor en comparación con los RNN (Pte, 2021).

1.2. Unidad Recurrente Cerrada (GRU)

GRU es una variante de LSTM que fue introducida por (Cho, y otros, 2014), la estructura interna de GRU es más simple y, por lo tanto, es más rápida de entrenar, ya que se necesitan menos cálculos para actualizar su estado oculto. GRU tiene dos puertas: puerta de reinicio y puerta de actualización La puerta de reinicio determina cómo combinar la nueva entrada con la celda de memoria anterior y la puerta de actualización define qué cantidad de la celda de memoria anterior se debe conservar (Manavi & Zhang, 2019).

2. Redes neuronales convolucionales (CNN)

CNN generalmente tiene un enfoque esencial para el reconocimiento de imágenes, sin embargo, se pueden utilizar en problemas de procesamiento del lenguaje natural. Las CNN son bastante diferentes de las redes neuronales de avance tradicionales, ya que tratan a la data como datos espaciales (Pte, 2021).

3. Redes Neuronales Transformers

Las redes Transformers, según su artículo de presentación realizado por Google en el año 2017, las redes Transformers están basadas en mecanismos de atención, dejando de lado la recurrencia en redes o convoluciones existentes, la relevancia de estas redes es que pueden llegar a ser paralelizables haciendo que el entrenamiento de estos modelos sea de mayor cantidad de datos en menor tiempo, las redes Transformers utilizan arquitecturas codificador y decodificador con múltiples capas de atención dentro de estos mecanismos, esto permite llegar a tener una mejor comprensión del contexto (Vaswani, y otros, 2017).

4.2.2. IA basada en lógica matemática

Según (McCarthy, 2007), “Lo que un programa sabe sobre el mundo en general, los hechos de la situación específica en la que debe actuar, y sus objetivos, están todos representados por oraciones de algún lenguaje lógico matemático. El programa decide qué hacer al inferir que ciertas acciones son apropiadas para lograr sus objetivos”.

La IA lógica implica representar el conocimiento del mundo de un agente, sus objetivos y la situación actual mediante oraciones en lógica. El agente decide qué hacer al inferir que cierta acción o curso de acción es apropiado para lograr los objetivos (McCarthy, Concepts of Logical AI, 2000).

4.2.3. Ingeniería de Búsqueda (Search)

Según (McCarthy, 2007), "Los programas de IA a menudo examinan un gran número de posibilidades, movimientos en un juego de ajedrez o inferencias mediante un programa de prueba de teoremas. Continuamente se hacen descubrimientos sobre cómo hacer esto de manera más eficiente en varios dominios".

4.2.4. Representación de Conocimiento

Según (McCarthy, 2007), "Los hechos sobre el mundo tienen que ser representados de alguna manera. Usualmente se utilizan lenguajes de lógica matemática".

4.2.5. Planning

Según (McCarthy, 2007), "Los programas de planificación comienzan con hechos generales sobre el mundo (especialmente hechos sobre los efectos de las acciones), hechos sobre la situación particular y una declaración de una meta. A partir de estos, generan una estrategia para lograr el objetivo. En los casos más comunes, la estrategia es solo una secuencia de acciones".

4.2.6. Heurísticas

Según (McCarthy, 2007), "Una heurística es una forma de tratar de descubrir algo o una idea incrustada en un programa. El término se utiliza de diversas formas en la IA. Las funciones heurísticas se utilizan en algunos enfoques de búsqueda para medir qué tan lejos parece estar un nodo en un árbol de búsqueda de un objetivo".

4.2.7. Programación Genética (Genes)

Según (McCarthy, 2007), "La programación genética es una técnica para hacer que los programas resuelvan una tarea emparejando programas Lisp aleatorios y seleccionando los más aptos en millones de generaciones".

El algoritmo genético es uno de los subconjuntos computacionales evolucionados que tienen una relación directa con el tema de la inteligencia artificial. Se le puede llamar un subconjunto de la inteligencia artificial. El algoritmo genético es un método de búsqueda que imita las reglas de la evolución biológica natural. El algoritmo genético

aplica la ley de supervivencia del más apto a una serie de soluciones de problemas para obtener mejores respuestas. En cada generación se obtienen mejores aproximaciones de las respuestas finales con la ayuda de un proceso de selección proporcional al valor de las respuestas y reproducción y utilizando operadores que imitan la genética natural. Este proceso adapta a las nuevas generaciones a las condiciones del problema (Manavi & Zhang, 2019).

4.3. **Ciencia de datos**

En la sección Ciencia de datos se expone una breve descripción sobre la metodología de análisis de datos CRISP-DM junto con las actividades que se realizan para esta metodología.

Metodología de análisis de datos CRISP-DM

Es un estándar de proceso de análisis de datos que describe los enfoques comúnmente utilizados para realizar el análisis de datos cuando el volumen de datos es alto. Es aplicable en diversas industrias (Nabati & Thoben, 2017).

CRISP-DM se describe en términos de un modelo de proceso jerárquico y cíclico compuesto por seis fases, cada una de las cuales consta de varias tareas genéricas, tareas especializadas e instancias de proceso. el modelo de proceso CRISP-DM creado por IBM, reporta un nivel de uso del 43% seguido por un 28% de empresas que utilizan su propia metodología. La metodología está diseñada para ser genérica, completa y estable, lo que significa que debe cubrir todo el proceso de desarrollo analítico para todas las aplicaciones posibles y debe ser válida para desarrollos no previstos (por ejemplo, nuevas técnicas de modelado analítico) (Kristoffersen, Aremu, Blomsma, Mikalef, & Li, 2019).

CRISP-DM se describe en términos de un modelo de proceso jerárquico y cíclico compuesto por seis fases, cada una de las cuales consta de varias tareas genéricas, tareas especializadas e instancias de proceso. el modelo de proceso CRISP-DM creado por IBM, reporta un nivel de uso del 43% seguido por un 28% de empresas que utilizan su propia metodología. La metodología está diseñada para ser genérica, completa y estable, lo que significa que debe cubrir todo el proceso de desarrollo analítico para todas las aplicaciones posibles y debe ser válida para desarrollos no previstos (por ejemplo, nuevas técnicas de modelado analítico) (Kristoffersen, Aremu, Blomsma, Mikalef, & Li, 2019).

Actividades de la Metodología CRISP-DM

A continuación, se explica las seis fases que existen dentro de la mitología CRISP-DM.

1. **Entendiendo el negocio**

La comprensión empresarial muestra los requisitos del problema de análisis de datos a través de requerimientos, supuestos, restricciones, riesgos y contingencias, costos y beneficios desde la perspectiva empresarial. Hace hincapié en comprender el objetivo del análisis de datos y reconocer muy bien los aspectos del problema (Nabati & Thoben, 2017).

2. **Recopilación y Comprensión de datos**

El proceso de comprensión de datos incluye la recopilación inicial de datos y familiarizarse con los datos, sus variables y dimensiones (Nabati & Thoben, 2017). Dentro de estas técnicas de extracción de datos se pueden tener distintas fuentes: bases de datos, imágenes, videos, web, redes sociales, IoT, etc.

Fuentes de Datos (Dataset)

2.1.IoT

IoT tiene como objetivo reunir todo bajo una única infraestructura de comunicación, dándonos el poder de controlar cualquier cosa desde cualquier lugar. Se están implementando poderosas herramientas de análisis de datos para recopilar y analizar datos de esos dispositivos conectados para hacer un cambio drástico en nuestra vida diaria. Es un concepto que tiene un enorme significado social, económico y técnico (Mridha, Hamid, & Asaduzzaman, 2019).

2.2.Redes Sociales

Las redes sociales y su uso se definen como aplicaciones de Internet que ayudan a los consumidores a compartir opiniones, conocimientos, experiencias y perspectivas (Kaplan & Haenlein, 2009). Las redes sociales se definen como un servicio basado en

la web que permite a las personas construir un perfil público o semippúblico dentro de un sistema delimitado, crear una lista de otros usuarios con los que comparten una conexión. Estos sitios permiten la interacción de usuarios con intereses similares a través del uso de Internet. (DiPietro, Crews, Gustafson, & Strick, 2012)

3. Preparación de datos

La preparación de datos contiene el filtrado, la agregación, la selección de parámetros y la construcción de un subconjunto de datos adecuado para el análisis (Nabati & Thoben, 2017).

Procesamiento del lenguaje natural

Uno de los pilares fundamentales de la inteligencia artificial es la manipulación de lenguaje natural mediante las herramientas de computación. Se basa en la utilización de lenguaje natural para proporcionar información de fácil procesamiento para la computadora, con la finalidad de facilitar el desarrollo de programas o modelos que ayudan a comprender el comportamiento humano (Cortez Vásquez, Vega Huerta, & Pariona Quispe, 2009).

4. Elección de modelo

Se ajustan diferentes modelos a los datos y se encuentran los valores óptimos. Los modelos relevantes cubiertos son el aprendizaje automático, la minería de datos y el análisis estadístico (Nabati & Thoben, 2017). En este apartado se definieran las técnicas del modelo, generación de diseño de pruebas, construcción del modelo y revisión de modelo.

4.1. Resumen de texto

El resumen automático de texto se clasifica en términos generales como extractivo o abstractivo. En el resumen extractivo, el algoritmo selecciona las líneas más importantes del párrafo de entrada y lo llama resumen. Estos son los sistemas que generan resúmenes copiando partes del documento mediante el despliegue de varias medidas de importancia y luego combinan esas partes para presentar un resumen. El

rango importante de la oración se basa en las características lingüísticas y estadísticas (Menaka, Thaker, Bhushan, & Karthik, 2020).

En el resumen abstracto, el algoritmo comprende el párrafo e intenta hacer un resumen novedoso usando frases y creando oraciones que no se encuentran en el pasaje de entrada. Estos son los sistemas que generan nuevas frases, ya sea reformulando o utilizando palabras que no estaban en el texto original (Menaka, Thaker, Bhushan, & Karthik, 2020).

Los métodos abstractivos suelen ser más concisos y parecidos a los humanos, mientras que los métodos extractivos no resumen completamente, sino que acortan el párrafo dado (Menaka, Thaker, Bhushan, & Karthik, 2020).

4.2. Análisis de datos

Análisis y Clasificación de sentimientos

El Análisis de sentimientos se encuentra en tres niveles de granularidad: nivel de documento, nivel de oración, y nivel de aspecto (Zhang, Wang, & Liu, 2018). Es la construcción de un modelo para recoger y categorizar las diferentes opiniones mediante la extracción del texto, posterior se identifica la información subjetiva de ciertas categorías, para predecir si el sentimiento es positivo, negativo o neutral, usando el proceso del lenguaje natural y las técnicas de aprendizaje automático (Paez & Monroy, 2020)

5. Modelado

En la etapa de evaluación se valora el modelo y el resultado obtenido de él, (Nabati & Thoben, 2017). Teniendo que revisar los criterios de éxito, medidas de calidad, revisión del proceso.

6. Despliegue o implementación

El conocimiento obtenido del modelado se discute con el usuario y se aplica al problema en acción (Nabati & Thoben, 2017). En este apartado se planifica el

despliegue, monitoreo y mantenimiento; además de debe generar un reporte final junto a una documentación de experiencia del proyecto.

4.4. **Calidad de servicio y satisfacción del cliente**

Según (CANTOS & MOLINER, 2001), "La Calidad de Servicio como el valor añadido que ofrecen, a un cliente cada vez más exigente, que busca en la actividad de consumo, aquel producto o servicio que puede satisfacer de una forma óptima sus necesidades, ofreciéndole, en definitiva, mejores resultados". Por lo tanto, para (CANTOS & MOLINER, 2001) "Un cliente satisfecho es un cliente leal que está dispuesto a realizar comentarios positivos a otros clientes. Así pues, tanto la Calidad de Servicio como la Satisfacción del cliente son críticas para entender la prosperidad de la organización".

Un buen servicio lleva consigo grandes consecuencias para la organización en todo aspecto. Desde dar a conocer su marca, producto o servicio, adquirir liderazgo en relación con la competencia, y sus stakeholders (PINEROS & BLANCO, 2009).

4.5. Metodología de desarrollo Scrum

En esta sección se aborda los temas relacionados hacia la metodología de desarrollo ágil Scrum. Se realiza la creación de ciclos de corto plazo para limitar el tiempo de desarrollo. Es altamente aceptado debido a que tiene la capacidad de introducir nuevos cambios o modificaciones con entregas en el menor tiempo posible.

Scrum cuenta con diferentes tipos de roles que están involucrados con el proyecto y el proceso anteriormente definido (Trigas Gallegos):

- **Product Owner**

  El Product Owner es responsable de maximizar el valor del producto resultante creado por el equipo de desarrollo. Este proceso puede variar ampliamente entre organizaciones, equipos Scrum e individuos. El Product Owner es la única persona responsable de gestionar la cartera de productos (Kelly, 2019).

- **Scrum Máster**

  Persona encargada de analizar y verificar que el modelo propuesto funcione correctamente sin presentar algún tipo de inconveniente. El Scrum Máster agrega valor al aceptar y actuar sobre esas responsabilidades. Enseña y entrena al propietario del producto y al equipo de desarrollo sobre cómo usar el marco Scrum correctamente. Más específicamente, la responsabilidad del Scrum Máster es asegurarse de que el propietario del producto y el equipo de desarrollo entiendan y cumplan con sus responsabilidades. El Scrum Máster es responsable de la capacidad del Scrum Team para funcionar como equipo (Fowler, 2018).

- **Scrum Team**

  Consiste en un grupo de personas que juntas crean, mantienen y mejoran un producto en particular. El equipo Scrum asume la responsabilidad del producto y realiza su trabajo sin necesidad de ayuda externa al equipo. Se forma inicialmente cuando se crea el producto por primera vez (Fowler, 2018).

Sprint

Según (Mariño & Alfonzo, 2014) los sprints, "son iteraciones de 1 a 4 semanas, y se suceden una detrás de otra. Al comienzo de cada Sprint, el equipo multifuncional selecciona los requisitos del cliente de una lista priorizada. Se comprometen a terminar los elementos al final del Sprint. Durante el Sprint no se pueden cambiar los elementos elegidos. Al final del Sprint, el equipo lo revisa con los interesados en el proyecto, y les enseña lo que han construido". Un Sprint consta de los siguientes eventos:

- **Planificación de Sprint**

  Se inicia cuando el equipo de trabajo establece el trabajo que se hará.

- **Scrum diario**

  En este evento se supervisa el progreso hacia al objetivo Sprint, se realiza a diario en reuniones que duran máximo 15 minutos.

- **Revisión del Sprint**

  El equipo Scrum presenta los resultados del Sprint para inspección, en base a esto se determinará que hacer a continuación, este evento dura máximo 4 horas y es una vez cada mes.

- **La retrospectiva del Sprint**

  El equipo scrum analiza los problemas que se tuvieron al ejecutar el Sprint y si esos problemas fueron resueltos o no, identificando así mejores que podrán ser útiles para aumentar la calidad y la eficacia.

5. Marco Metodológico

Dentro de esta sección se observa la propuesta de solución planteada para lograr los objetivos del proyecto, mediante fases. También se explica la metodología que utiliza dentro del proyecto.

5.1. Propuesta de Solución

El siguiente gráfico se presenta el modelo de la estructura de la aplicación híbrida, y posterior se explican las fases del proyecto.

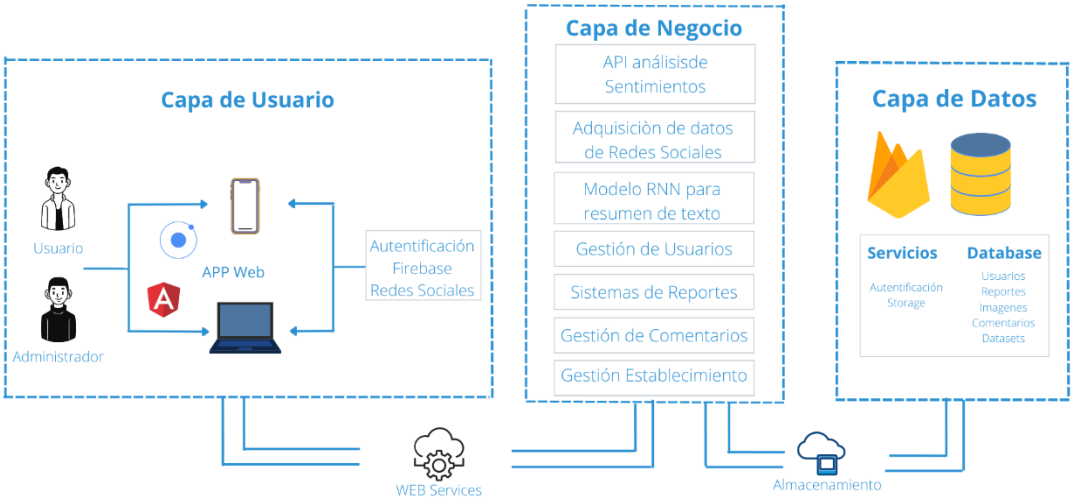


Figura 1. Estructura de la Aplicación

El presente proyecto se divide en 5 fases:

I. En la **primera fase** se estudia los fundamentos de Deep Learning, redes neuronales, procesamiento de lenguaje natural, IONIC, Angular, Django y herramientas de extracción de datos, con el fin de tener de claros los conceptos que se aplican en cada uno de los módulos de la aplicación híbrida inteligente.

II. En la **segunda fase** está enfocada al diseño y desarrollo del modelo de análisis de datos en fases para resumir y representar comentarios.

Como primer paso se realiza la importación de varias librerías que permiten trabajar con texto, como Keras, Tensorflow, pandas y Torch entre las más destacadas. Posterior se procede a cargar el dataset con los comentarios acerca de

bares y restaurantes. Se procede a realizar la limpieza y preprocesamiento sobre estos datos. Con el conjunto de datos limpio se procede a tokenizar el dataset, esto permite generar el modelo de red neuronal Transformer, para el entrenamiento y testing del algoritmo, todos estos están bajo el Transformer T5 desarrollado por Google, la optimización del modelo estará en función a los resultados obtenidos. A continuación, se verifica que tan bueno es el algoritmo mediante medidas de calidad a través de ROUGE. Finalmente, se procede a la representación del resumen del comentario original del usuario.

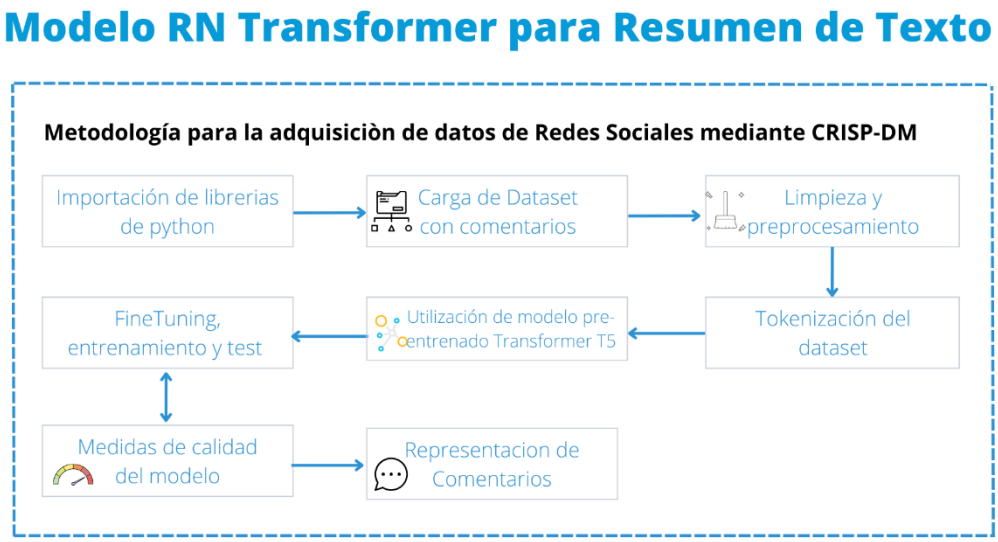


Figura 2. Proceso de Red Neuronal Transformers para resumen de texto

III. La **tercera fase** consiste en el diseño y desarrollar de la aplicación híbrida y el módulo de adquisición de datos de redes sociales.

Para el diseño y desarrollo de la aplicación híbrida, se tiene dos tipos de usuarios (Usuario, Administrador), estos se conectan con el sistema a través de la aplicación híbrida desarrollada en framework Angular junto a IONIC. Se exige una autentificación por parte del usuario mediante Firebase permitiendo el ingreso a cada uno de los módulos.

Los módulos propuestos para el usuario son la gestión de comentarios, el sujeto dispone de una interfaz gráfica para ingresar el comentario acerca del bar y restaurantes, esta información es almacenada en la base de datos para su posterior manipulación.

Los módulos propuestos para el administrador son sistemas de reportes y gestión de establecimiento. En el primer módulo, el sujeto dispone de una interfaz gráfica para ver la representación de los comentarios, que son representados a través de

diagramas, información, opiniones más relevantes, acerca del establecimiento y el segundo módulo se enfoca a mostrar información básica del establecimiento.

Por otro lado, tenemos al modelo de adquisición de datos a través de redes sociales, se eligió Facebook como ejemplo. Para acceder a la API de Facebook, se crea una cuenta que debe ser aprobar por Facebook, a fin de obtener un token de acceso para la realización de consultas. El presente proyecto está centrado en la recolección de comentarios de una o varias publicaciones del negocio, para proceder a limpiar y procesar los datos. Posterior a esto, se realiza un correcto mapeo y filtrado de comentarios para generar un dataset de comentarios que finalmente se almacenan en la base de datos en Firebase.

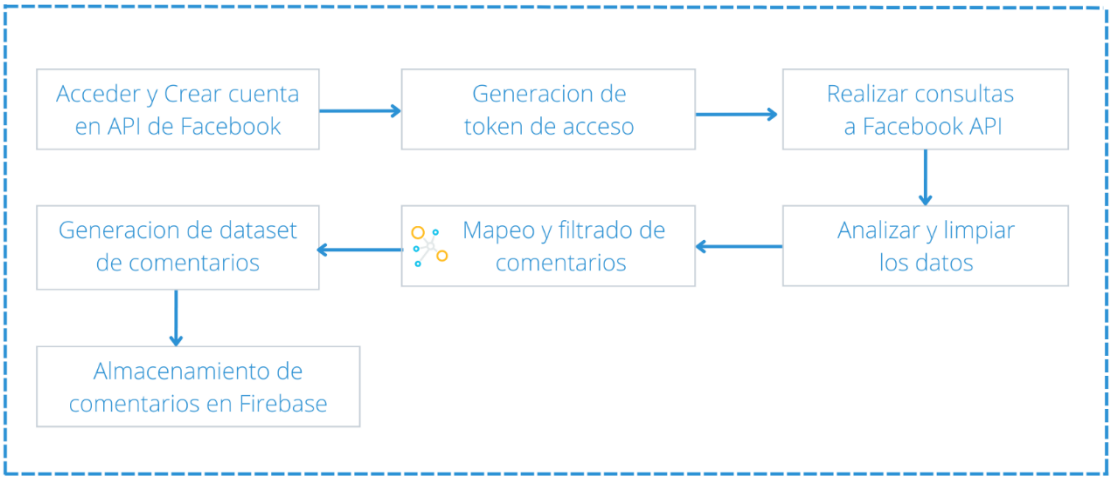


Figura 3. Modelo de adquisición de datos a través de Redes Sociales

IV. En la **cuarte fase** se tiene el integrar el modelo de datos como servicio dentro de la aplicación hibrida.

En esta fase, administrador y usuario, hacen uso del modelo de sistemas de reportes, el usuario para el almacenamiento de comentarios y el administrador para la representación de dichos comentarios.

Cuando el usuario registra un nuevo comentario, se envía a través de web services como un método post con destino a nuestro backend que tiene el modelo para resumen de comentarios, en este se recibe el comentario, se procesa y se obtiene su resumen, posteriormente se almacena en la base de datos, con su respectivo sentimiento, para crear la representación de los datos. Mientras tanto, el usuario visualiza en la interfaz que agregó un nuevo comentario.

Cuando el Administrador ingresa a la interfaz de reportes de nuestra aplicación, el sistema es el responsable de cargar cada tipo de representación del resumen generado mediante gráficas de barras, pastel, estadística descriptiva, tablas de información acerca de comentarios, según el tipo de comentario, palabras más repetidas enfocadas al negocio como platos más mencionados de manera positiva, negativa o neutral. Toda la información será recuperada de la base de datos de Firebase. Al abrir la interfaz de reporte, se realiza una petición get a través de web services con destino a nuestro backend en donde se encuentra el modelo para representación de comentarios, aquí se generan las diferentes formas de representación de comentarios y se almacenan en nuestro base de datos.

V. La **quinta fase** y última del proyecto es diseñar un plan de experimentación o también conocido como un plan de pruebas del modelo de resumen de texto, modelo de adquisición de datos de redes sociales y del funcionamiento de aplicación híbrida inteligente, con la finalidad de calificar y verificar el correcto funcionamiento del software.

5.2. Metodología

Se abordan dos secciones de metodología, una enfocada al análisis de datos y la segunda al desarrollo de software para la aplicación hibrida.

5.2.1. Modelo de Análisis aplicando metodología CRISP-DM

En las siguientes secciones se aborda como se aplicó la metodología CRISP-DM al proyecto.

A. Entendiendo el negocio

Adquisición de requerimientos al cliente, dentro de este proyecto la finalidad es obtener un resumen de un comentario, ya sea ingresado a través de la aplicación móvil o también analizar los comentarios de una red social del cliente (Facebook), con el fin de generar reportes que brinden conocimiento al negocio para y brindar ayuda en la toma de decisiones.

B. Recopilación y Comprensión de datos

Para este proceso se trabajó con varios datasets públicos en idioma en inglés que hacían referencia a comentarios de productos de Amazon, también se nos proporcionó un dataset de restaurantes con comentarios en español, además se obtuvo comentarios de publicaciones en redes sociales, ya sea mediante API de la misma aplicación, o realizando raspado web (web scraping). Al final se obtiene un conjunto de datos con los cuales se puede trabajar tanto en inglés como en español.

C. Preparación de datos

Una vez obtenida esta información, se pasa al proceso de limpieza y preprocesamiento para obtener un dataset limpio y categorizado, dentro de este proceso de limpieza se hace uso de técnicas de procesamiento del lenguaje natural, en donde se eliminan valores nulos, caracteres especiales, pasar texto a minúsculas, quitar espacios extra o tabulaciones, eliminación de etiquetas HTML, mapeo de contracciones en idioma inglés, eliminación de apóstrofes ('s), eliminar cualquier texto dentro de paréntesis, eliminar palabras cortas, esta limpieza se realiza dentro de los comentarios y su correspondiente resumen; este dataset se utiliza para posterior entrenamiento, sin embargo, para realizar EDA (Análisis Exploratorio de datos), se procede a eliminar stop words, crear una bolsa de palabras, generación de nube de palabras, palabras con mayor frecuencia, ngramas y distribución de secuencias de oraciones para resumen y texto, este último nos dará una perspectiva de la distribución de longitud promedio que tendrán el comentario y resumen.

D. Elección de modelo

Dentro de este proyecto se utiliza el modelo de red neuronal transformer de transferencia de texto a texto T5 publicado por (Raffel, y otros, 2020), detrás de este modelo se plantea una arquitecta codificador – decodificador en las cuales se adjuntan mecanismos de atención.

Según (Etemad, Abidi, & Chhabra, 2021), el Codificador, contiene dos subcapas, la capa de autoatención (self – attention layer) y una capa de avance (feed – forward layer); como primer paso el codificador es el encargado de codificar el texto de entrada en un vector mediante algoritmos de incrustación (embedding algorithms), posterior se aplica codificación posicional para mantener la secuencia de tokens; con esta secuencia vamos a pasar a través de la capa de autoatención (self attention layer).

En esta capa, cada palabra contiene un valor de atención individual, el cual es calculado utilizando la llave, la consulta y el valor del vector obtenido aplicando producto escalar de vectores que fueron inicializados aleatoriamente. Después de calcular el valor de atención para cada palabra, todos estos valores se concatenan, con la finalidad de calcular un vector de atención final. La salida de capa de atención (attention layer) pasará a través de una capa de avance (feed – forward). Este proceso se repite en las próximas seis capas del codificador. Una vez que se llega al último codificador, se obtiene una salida, esta será enviada al decodificador.

Según (Etemad, Abidi, & Chhabra, 2021), el decodificador, cada decodificador asignado contiene 3 subcapas: encoder-decoder attention, feed-forward attention y multi-head attention. Con la entrada recibida del codificador, este módulo procede a aplicar encoder – decorder attention generando una salida. La capa de multi – head nos proporciona que el modelo se enfoque en la parte apropiada de las oraciones. Este proceso se repite hasta que se llega al token de salida. Posterior, el decodificador envía cada salida generada a la parte más inferior del decodificador. En el modelo original del transformador este cuenta con seis capas de codificador y seis de decodificador.

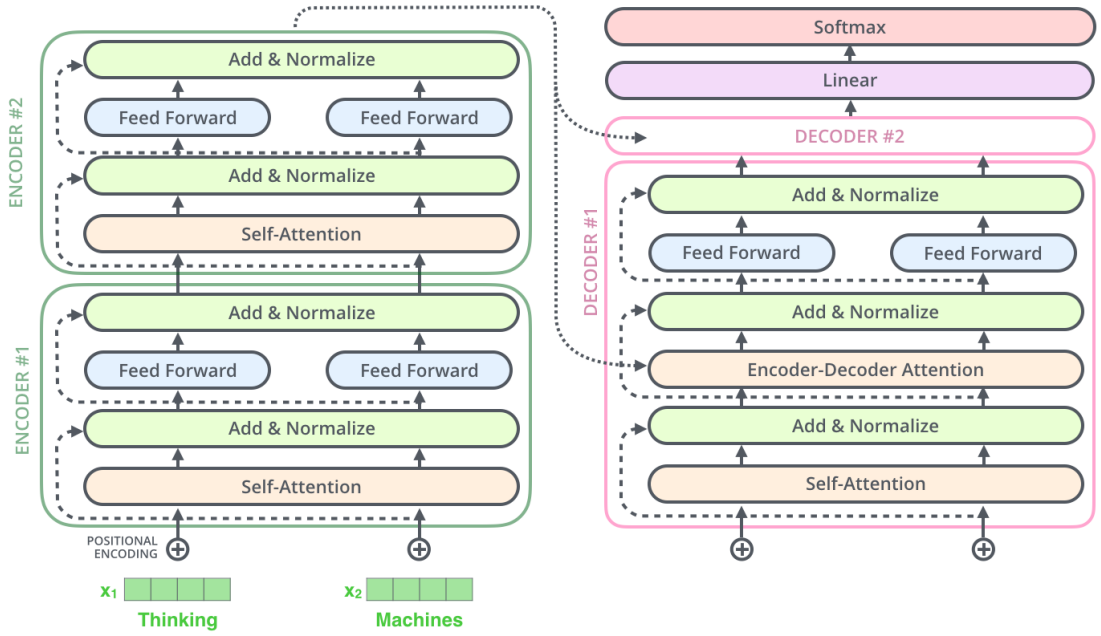


Figura 4. Estructura del modelo Transformers T5 realizado por (Alammar, 2018)

Para este proyecto se toma como modelo base a Transferencia de texto a texto (T5), que tiene como entrada y salida al texto, este modelo contiene las siguientes características, traducción automática, tarea de clasificación, tarea de regresión, resumen de documentos y otras tareas de secuencia a secuencia. Para esta parte se realiza un aprendizaje por transferencia, debido a que el modelo es entrenado con un conjunto de datos abundantemente disponible. Según (Garg, y otros, 2021) muestra que T5 tiene mejores resultados en el resumen de texto en comparación con el modelo BART, sin embargo

(Deokar & Shah, 2021) nos dice que Bart se comporta mejor que T5, pero debido a la documentación que existe alrededor de T5 y varios ejemplos para el objetivo que necesitamos, se enfocó el proyecto a este modelo.

T5 es un modelo pre - entrenado con texto extraído de la web, con 750 gigabytes de texto en inglés, este conjunto de datos se encuentra abierto al público dentro de la página oficial de TensorFlow c4. Todos los archivos de configuración han sido publicados en internet, así como sus modelos pre – entrenados (Raffel, y otros, 2020). Se encuentran liberados 5 variantes del modelo: modelo pequeño, base, grande, 3 Billones y 11 Billones, los modelos funcionan muy cerca del nivel humano, en el paper se crea una métrica para la predicción de esto.

Parámetros del modelo.

- Para el proyecto se trabaja con el modelo base de T5 (t5-base)
- En la tokenización se hace uso de un modelo pre – entrenado (t5-base) tokenizador
- El conjunto de Datos con los que se realiza el entrenamiento está en idioma español, con doscientos registros, en donde 180 fueron para entrenamiento y 20 se enfocaron en testing.
- Dentro del entrenamiento se configuró los siguientes parámetros, conjunto de entrenamiento y de prueba, una longitud máxima de token para texto y resumen (200 y 18 respectivamente), esta información viene dado por un análisis previo de la longitud del conjunto de entrenamiento y de prueba. Para el tamaño de batch se definió en 12 y una cantidad de epochs en 8, early stopping se encuentra deshabilitado y optimizador Adam, estos parámetros pueden ser mayores, si se tiene un buen equipo para entrenar nuestro modelo.

5.2.2. Metodología SCRUM

Para el desarrollo del proyecto se utiliza la metodología SCRUM, y se definen 4 sprints con sus respectivas actividades. Se aclara que por cada finalización de un sprint se procede a realizar el siguiente, para tener control sobre del desarrollo del proyecto desde el principio hasta el final.

Dentro de esta metodología, el Product Owner es el encargado de listar los requerimientos de la aplicación híbrida inteligente, para detallar cada uno de los sprints, acorde al nivel de importancia y necesidad dentro del proyecto. Scrum Máster es el sujeto encargado de conformar el equipo de trabajo, delegar las tareas que deben cumplir durante los sprints, ejecutar reuniones una o dos veces por semana, para solventar dudas o percances generados en el desarrollo de la tarea asignada. Cada Sprint tiene su respectiva revisión y se redefinirán la prioridad de las tareas con el fin de avanzar de manera más ágil el proyecto. Generalmente, si el Sprint dura alrededor de 4 semanas se estima que el Sprint Review es de 8 horas.

Roles Scrum:

Product Owner: Ing. Remigio Hurtado (RH).

Scrum Máster: Katherine Michelle Barrera Barrera (KB).

Scrum Team:

Developer N°1: Juan Carlos Barrera Barrera (JB).

Developer N°2: Katherine Michelle Barrera Barrera (KB).

OE1. Estudiar y conocer los fundamentos de Deep Learning, procesamiento de lenguaje natural, IONIC, Angular, Django y herramientas de extracción de datos.

No.	Actividad
1.	Estudiar los fundamentos de Deep Learning
2.	Estudiar los fundamentos de procesamiento de lenguaje natural
3.	Estudiar herramientas para extracción de datos
4.	Estudiar los fundamentos de Framework para desarrollo móvil y web como IONIC y Angular.
ER.	Escritura y revisión de evidencias

OE2. Diseñar y desarrollar el modelo de análisis de datos en fases para resumir y representar comentarios.

No.	Actividad
1.	Proceso de recolección de datos
2.	Preparación y limpieza de datos
3.	Diseño y desarrollo del modelo de análisis de datos para resumir comentarios
4.	Análisis de resultados obtenidos del modelo aplicado
5.	Diseño y desarrollo para representación de datos
ER.	Escritura y revisión de evidencias

OE3. Diseñar y desarrollar la aplicación híbrida y el módulo de adquisición de datos de redes sociales.

No.	Actividad
1.	Especificación de requerimientos funcionales y no funcionales.
2.	Diseño y desarrollo del módulo Usuario
3.	Diseño y desarrollo del módulo Administrador
4.	Diseño y desarrollo del módulo de Reportes
5.	Diseño y desarrollo del módulo de Comentarios
6.	Diseño y desarrollo del módulo de Establecimiento
7.	Desarrollo del módulo Autentificación de usuario a través de Firebase
8.	Integración del módulo de adquisición de datos de redes sociales.
9.	Diseño y desarrollo de la base de datos dentro de Firebase
ER.	Escritura y revisión de evidencias

OE4. Integrar el modelo de datos como servicio dentro de la aplicación hibrida.

No.	Actividad
1.	Integración del servicio de API REST para agregar nuevo comentario realizado por el usuario.
2.	Integración del servicio de API REST para el despliegue del reporte para el usuario administrador.
ER.	Escritura y revisión de evidencias

OE5. Diseñar un plan de experimentación del modelo de análisis y de la aplicación híbrida

No.	Actividad
1.	Diseño un plan de pruebas funcionales y no funcionales
2.	Diseño y desarrollo de pruebas para el módulo Usuario
3.	Diseño y desarrollo de pruebas para el módulo Administrador
4.	Diseño y desarrollo de pruebas para el módulo de Reportes
5.	Diseño y desarrollo de pruebas para el módulo de Comentarios
6.	Diseño y desarrollo de pruebas para el módulo de Establecimiento
7.	Diseño y desarrollo de pruebas para el módulo Autentificación
ER.	Escritura y revisión de evidencias

A continuación, se presentan las actividades de cada objetivo separados en cinco tablas que representan un sprint con el fin de cumplir con la metodología Scrum.

Sprint 1	
OE.3	
	ACT.1. Especificación de requerimientos funcionales y no funcionales.
OE.1	
	ACT.1. Estudio de los fundamentos de Deep Learning
	ACT.2. Estudio de los fundamentos de procesamiento de lenguaje natural (PLN)
	ACT.3. Estudio de herramientas para extracción de datos
OE.2	
	ACT.1. Proceso de recolección de datos
	ACT.2. Preparación y limpieza de datos
	ACT.3. Diseño y desarrollo del modelo aplicado
	ACT.4. Diseño y desarrollo para representación de datos
OE.1. **OE.2.** **OE.3.**	
	ACT.ER. Escritura y revisión de evidencias

Sprint 2	
OE.1	
	ACT.4. Estudio de los fundamentos de Framework para desarrollo móvil y web como IONIC y Angular.
OE.3	
	ACT.2. Diseño y desarrollo del módulo Usuario
	ACT.3. Diseño y desarrollo del módulo Administrador
	ACT.4. Diseño y desarrollo del módulo de Reportes
	ACT.5. Diseño y desarrollo del módulo de Comentarios
	ACT.6. Diseño y desarrollo del módulo de Establecimiento
	ACT.7. Desarrollo del módulo Autentificación de usuario a través de Firebase
	ACT.8. Integración del módulo de adquisición de datos de redes sociales
	ACT.9. Diseño y desarrollo de la base de datos dentro de Firebase
OE.1. **OE.3.**	
	ACT.ER. Escritura y revisión de evidencias

Sprint 3	
OE. 4	
	ACT.1. Integración del servicio de API REST para agregar nuevo comentario realizado por el usuario
	ACT.2. Integración del servicio de API REST para el despliegue del reporte para el usuario administrador
OE.4.	
	ACT.ER. Escritura y revisión de evidencias

Sprint 4	
OE. 5	
	ACT.1. Diseño y desarrollo de plan de pruebas funcionales y no funcionales
	ACT.2. Diseño y desarrollo de pruebas para el módulo Usuario
	ACT.3. Diseño y desarrollo de pruebas para el módulo Administrador
	ACT.4. Diseño y desarrollo de pruebas para el módulo de Reportes
	ACT.5. Diseño y desarrollo de pruebas para el módulo de Comentarios
	ACT.6. Diseño y desarrollo de pruebas para el módulo de Establecimiento
	ACT.7. Diseño y desarrollo de pruebas para el módulo Autentificación
	ACT.8. Diseño y desarrollo de informe
OE.5.	
	ACT.ER. Escritura y revisión de evidencias

6. Resultados

En esta sección, se presentarán los resultados obtenidos del desarrollo de la aplicación, como prime punto se presenta los requerimientos funcionales, no funcionales y los prototipos realizados para la aplicación. Posterior a ello se presenta un diseño de alto nivel que contiene vista de desarrollo y física, en la siguiente sección se presenta el diseño detalle con un diagrama de actividades y clases. Posterior a ello se muestra el resultado del desarrollo del modelo de análisis separado en dos fases que son modelado e implementación y despliegue como servicio. La siguiente sección presenta la codificación de la aplicación híbrida inteligente y finalmente se presenta las pruebas unitarias realizadas con base en los requerimientos funcionales.

6.1. Especificación de Requerimientos

En esta sección se plantean los requerimientos funcionales y no funcionales, que se obtuvo de la entrevista realizada con el Product Owner.

6.1.1. Requerimientos Funcionales

Tabla 1. Requerimiento para inicio de sesión

Nombre	**Iniciar sesión de usuarios**	**Identificador**	**RF-01**
		Prioridad	Media
		Servicio	Módulo de gestión de usuarios
Descripción	Función para iniciar sesión para usuarios dentro de la aplicación. Existen dos roles, de administrador y usuario.		
Entrada	Ingreso de correo/contraseña o puede iniciar sesión mediante Google		
Proceso	1. Ingreso de correo 2. Ingreso de contraseña 3. Verificación de datos 4. Verifico la existencia del usuario 5. Acepto la autenticación		
Salida	Token de autentificación validado para navegar dentro de la interfaz según el rol.		

Tabla 2. Requerimiento para registrar usuarios

Nombre	**Registrar usuario/cliente**	**Identificador**	**RF-02**
		Prioridad	Media
		Servicio	Módulo de gestión de usuario cliente
Descripción	Función para registrar usuarios en la aplicación, solo para el rol de usuario. El administrador será creado desde la base de datos		
Entrada	Ingreso de metadatos en la interfaz de registro		
Proceso	1. Ingresar datos 2. Validar datos 3. Guardar datos 4. Crear token de autentificación para nuevo usuario		
Salida	Obtener token de autentificación		

Tabla 3. Requerimiento para la creación de usuario/cliente

Nombre	Creación de comentario del usuario/cliente	Identificador	RF-03
		Prioridad	Alta
		Servicio	Módulo de gestión de usuario cliente
Descripción	Función para crear un comentario con un usuario autentificado		
Entrada	Usuario autentificado Ingresar comentario		
Proceso	1. Tener usuario autentificado 2. Obtener comentario del usuario 3. Se clasifica en la API externa para clasificación de sentimientos 4. Hacemos uso API REST de Django para guardar el comentario de Firebase		
Salida	Comentario guardado en la base de datos Mostrar notificación al usuario		

Tabla 4. Requerimiento para listar comentarios cliente/usuario

Nombre	Listar comentarios del usuario/cliente	Identificador	RF-04
		Prioridad	Baja
		Servicio	Módulo de gestión de usuario cliente
Descripción	Función para que el usuario pueda listar los comentarios que ha realizado en la aplicación		
Entrada	El usuario debe estar autentificado, para obtener su token		
Proceso	1. Tener el usuario autentificado 2. Llamar al método de listar por usuario 3. Consulta a la Base de datos 4. Obtener resultado de consulta 5. Mostrar lista de comentarios al usuario		
Salida	Lista de comentarios de un usuario específico		

Tabla 5. Requerimiento para editar información básica del usuario

Nombre	Editar información básica del usuario	Identificador	RF-05
		Prioridad	Baja
		Servicio	Módulo de gestión de usuario cliente
Descripción	Función para editar la información personal del usuario y modificar en la base de datos.		
Entrada	1. Usuario autentificado 2. Datos personales modificados		
Proceso	1. Usuario autentificado 2. Usuario ingresa datos personales modificados 3. Validar datos modificados 4. Remplazar información personal del usuario en la base de datos		
Salida	Datos personales del usuario modificados		

Tabla 6. Requerimiento para listar comentarios del administrador

Nombre	Lista de comentarios para administrador	Identificador	RF-06
		Prioridad	Alta
		Servicio	Módulo de gestión de usuario administrador
Descripción	Función para listar todos los comentarios para el administrador		
Entrada	Administrador autentificado		
Proceso	1. Usuario administrador autentificado 2. Llamar método listar comentarios 3. Obtener lista de comentarios 4. Mostrar lista de comentarios		
Salida	Lista completa de cometarios		

Tabla 7. Requerimientos para generar reportes de comentarios

Nombre	Reporte de comentarios para administrador	Identificador	GU-07
		Prioridad	Alta
		Servicio	Módulo de gestión de usuario administrador
Descripción	Función para mostrar mediante gráficas como diagramas de barras, nubes de palabras, entre otros, información relevante obtenida de los comentarios		
Entrada	Usuario administrador autentificado		
Proceso	1. Administrador autentificado 2. Llamar a la función reportes 3. Enviar consultas hacia API REST 4. Obtener datos para gráficos 5. Crear gráficas con la información actualizada 6. Mostrar graficas del reporte al usuario		
Salida	Reporte con graficas para el usuario		

Tabla 8. Requerimiento para recolección de datos de redes sociales

Nombre	Recolección de datos de redes sociales	Identificador	RF-08
		Prioridad	Alta
		Servicio	Módulo de recolección de datos de redes sociales
Descripción	Función para recolectar comentarios mediante API de Facebook		
Entrada	Tener una página publica en Facebook		
Proceso	1. Obtener token de Facebook 2. Crear una aplicación en Graph Api 3. Crear una publicación 4. Obtener id de publicación y página 5. Llamada al método de obtención de comentarios 6. Almacenamiento de comentarios en archivo csv		
Salida	Archivo csv con los comentarios		

Tabla 9. Requerimiento para la creación del modelo para resumen de comentarios

Nombre	Creación del modelo para resumen de comentarios	Identificador	RF-09
		Prioridad	Alta
		Servicio	Módulo modelo de resumen de comentarios
Descripción	Un modelo de análisis para obtener el resumen de un comentario obtenido mediante la aplicación propia o mediante la red social de Facebook		
Entrada	Dataset de comentarios		
Proceso	1. Limpieza de datos 2. Generación modelo 3. Representación de comentarios		
Salida	Resumen de un comentario		

Tabla 10. Requerimiento para preparación de datos

Nombre	Preparación de datos	Identificador	RF-10
		Prioridad	Alta
		Servicio	Módulo modelo de resumen de comentarios
Descripción	Cuaderno de Jypiter para la preparación de datos		
Entrada	Dataset Original		
Proceso	1. Limpieza de datos 2. Tratamiento inicial de variables 3. Trasformaciones 4. Reducción		
Salida	1. Archivo csv con el dataset procesado 2. Pickle transformador		

Tabla 11. Requerimiento para creación del modelo de resumen de texto

Nombre	Creación de modelo para resumen de texto	Identificador	RF-11
		Prioridad	Alta
		Servicio	Módulo modelo de resumen de comentarios
Descripción	Cuaderno de Jypiter para la creación del modelo para resumen de texto		
Entrada	1. Archivo csv con el dataset procesado 2. Pickle transformador		
Proceso	1. Entrenamiento 2. Optimización 3. Evaluación con Test		
Salida	Objeto del modelo (arquitectura y pesos) Checkpoint		

Tabla 12. Requerimiento para predicción de nuevos valores

Nombre	Predicción con nuevos samples	Identificador	RF-12
		Prioridad	Alta
		Servicio	Módulo modelo de resumen de comentarios
Descripción	Cuaderno de Jypiter para la predicción con nuevos samples		
Entrada	1. Pickle transformador 2. Objeto del Modelo		
Proceso	1. Preparación (limpieza de un nuevo comentario) 2. Predicción		
Salida	Resumen de comentario nuevo		

6.1.2. Requerimientos No Funcionales

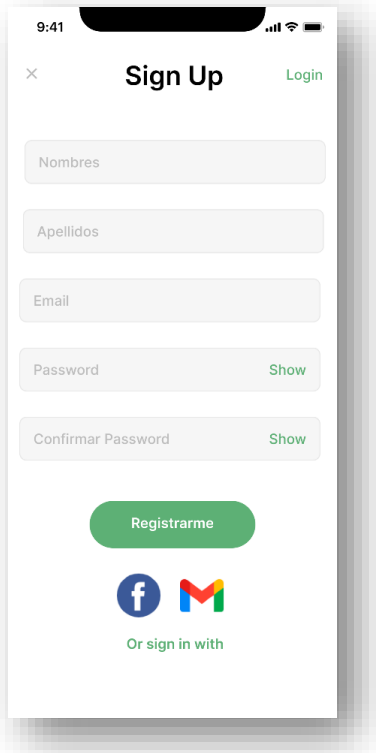
Tabla 13. Requerimientos no funcionales para la aplicación

RNF	
Eficiencia	La aplicación debe responder todo tipo de consulta en un rango menor a 6 segundos.
	Los datos modificados en la base de datos deben ser actualizados en un rango menor a 5 segundos.
Seguridad lógica y de datos	Los permisos de acceso serán modificados solamente por el administrador de acceso a datos.
	Se usa autenticación para el acceso de usuarios a la aplicación
	Se controla la contraseña del usuario, siendo mayor a 8 dígitos
	Se crean los permisos dentro de la aplicación para usuario administrador y usuario cliente.
Usabilidad	El aprendizaje para el manejo de aplicación por parte de los usuarios debe ser menor a 30 minutos
	Se mostrará mensajes informativos que ayuden al usuario a conocer el error que está cometiendo al momento de ingresar credenciales
	La aplicación debe tener un diseño Responsive, de forma que sea correctamente visible en múltiples dispositivos.
	Las interfaces deben ser amables con el usuario y proporcionar una fácil interacción
	La aplicación tendrá un modelo minimalista que ayudará hacer fácil su uso.
	Se debe manejar un menú para facilitar la navegación para el usuario final

6.1.3. Diseño de Wireframes o Prototipos

En este apartado se presentan los prototipos que se pretenden realizar para la aplicación.

Módulo de gestión de usuarios, iniciar sesión, crear cuenta y agregar información adicional de usuario.



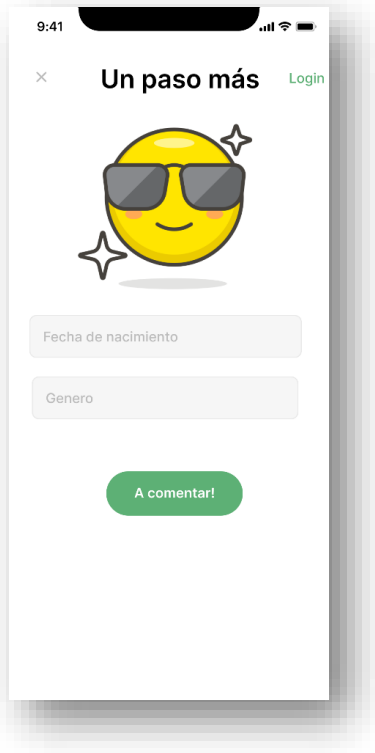
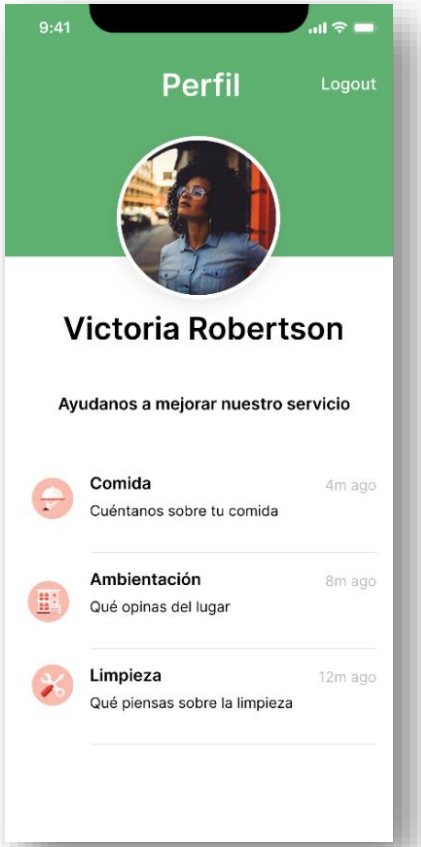
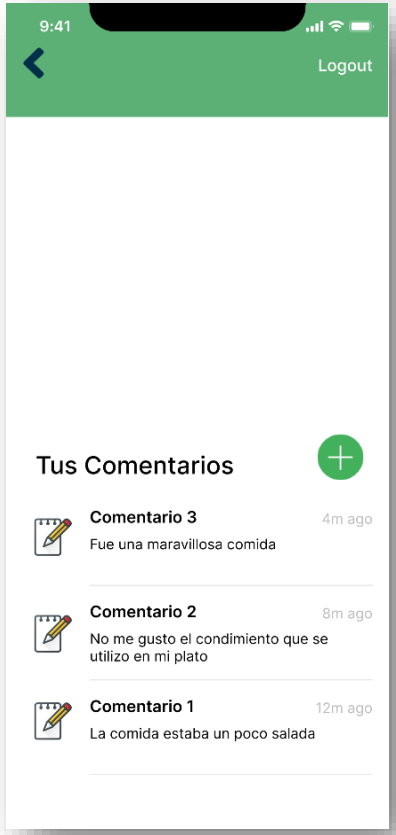


Figura 5. Inicio de Sesión	Figura 6. Registro de Usuarios	Figura 7. Ingreso de Datos adicionales

Módulo de gestión de usuario, página para perfil de usuario con sus comentarios, añadir nuevo comentario y listar su comentario





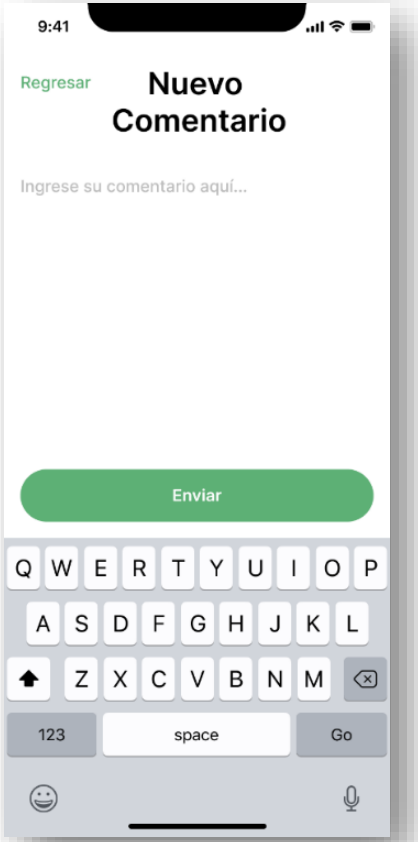


Figura 8. Perfil de Usuario	Figura 9. Lista de Comentarios	Figura 10. Agregar nuevo comentario

Módulo de gestión de usuario - administrador, página con frecuencia de palabras, nube de palabras y página para listar el resumen de todos los usuarios

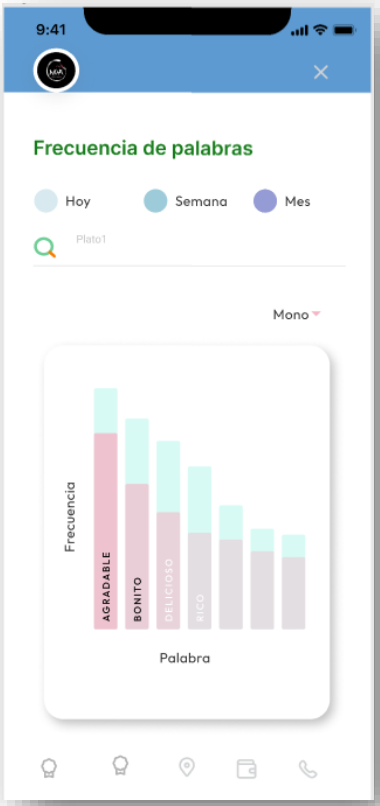


Figura 11. Pagina frecuencia de palabras

Figura 12. Nube de palabras

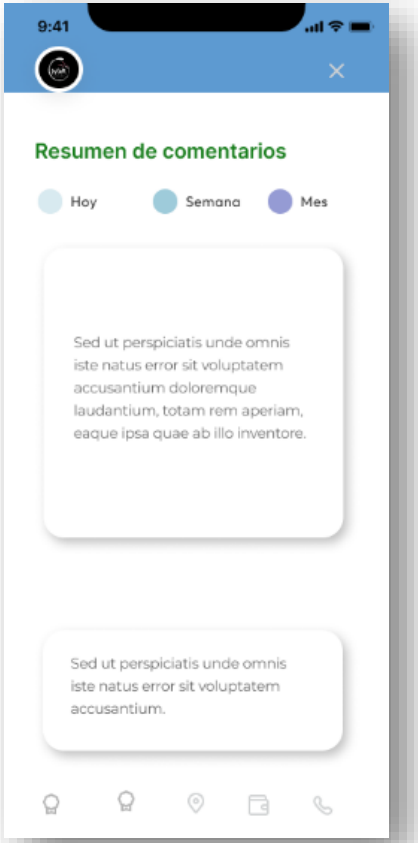


Figura 13. Pagina para listar el resumen de comentarios

6.2. Diseño de Alto Nivel

En el diseño de alto nivel se pretende mostrar cómo se ha implementado la solución del problema, mediante las vistas de desarrollo, física y una arquitectura API.

6.2.1. Vista de Desarrollo

En la vista de desarrollo, con base a lo que está planteado en la solución del problema, se ha diseñado una arquitectura de n capas, para esto en el proyecto se plantea 3 capas: capa de usuario, negocio y finalmente de datos. Dentro de cada capa, los componentes existentes cumplen funciones específicas.

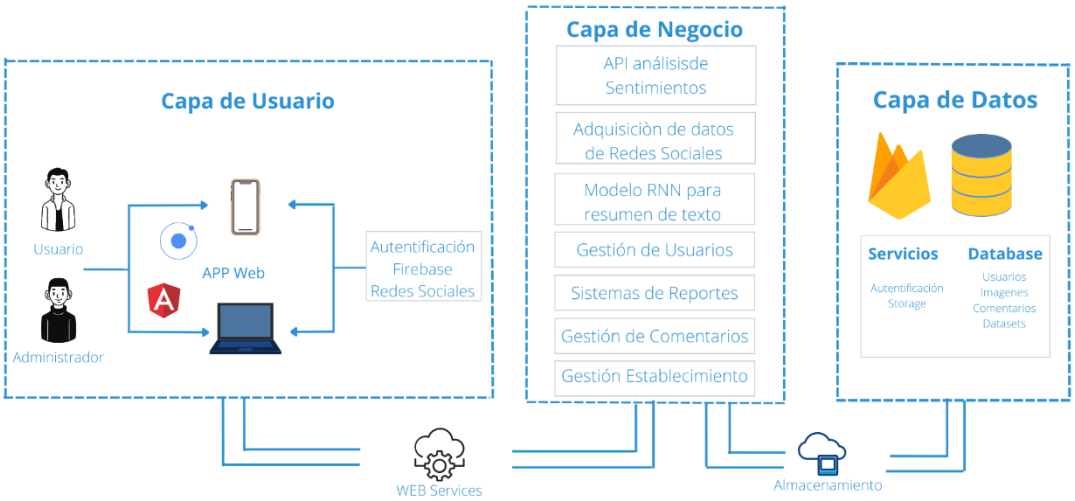


Figura 14. Arquitectura física de la aplicación hibrida

En primer lugar, tenemos a la capa de datos, la responsable de almacenar la información acerca de los usuarios, imágenes que se generan, comentarios tanto de la aplicación desarrollado como de las redes sociales obtenidas (Facebook), además de esto en esta capa se utilizan los recursos de autentificación y storage, que servirán para el ingreso a la aplicación tanto de los usuarios como del administrador.

En segundo lugar, se tiene la capa de negocio, esta capa hará de API Rest, permitiendo conectar a la capa de usuario con la de datos, está capa esta desarrollada bajo el framework de Django, en donde se tiene varias rutas a los distintos servicios implementados, gestión de comentarios, de reportes, de redes sociales, entre otros. Esta capa se conecta a una API pública que nos permite realizar análisis de sentimientos.

Por último, tenemos la capa de presentación, la cual está desarrollada bajo el framework de IONIC, en donde existe la gestión de usuarios, para el proyecto se plantea dos usuarios: usuario normal y administrador. El usuario normal podrá tener acceso al módulo de gestión de comentarios, permitiendo ver información del negocio, premios que puede obtener por participar en la aplicación y como aportar al establecimiento e ingresar un comentario sobre el mismo, por otro lado, tenemos al administrador el cual tiene acceso al módulo de gestión de reportes, en donde se mostraran gráficas obtenidas de los comentarios extraídos ya sea de la aplicación o de redes sociales, estas gráficas están relacionadas con las edades de los usuarios, sus opiniones, platos que más gusta a la gente, que opinan sus clientes del establecimiento, entre otros.

6.2.2. Vista física

A través de este diagrama se puede entender de una mejor para cómo se comunica los diferentes componentes del sistema, para esto es necesario un diagrama de despliegue, en la siguiente figura se aprecia como se encuentran estos módulos o componentes.

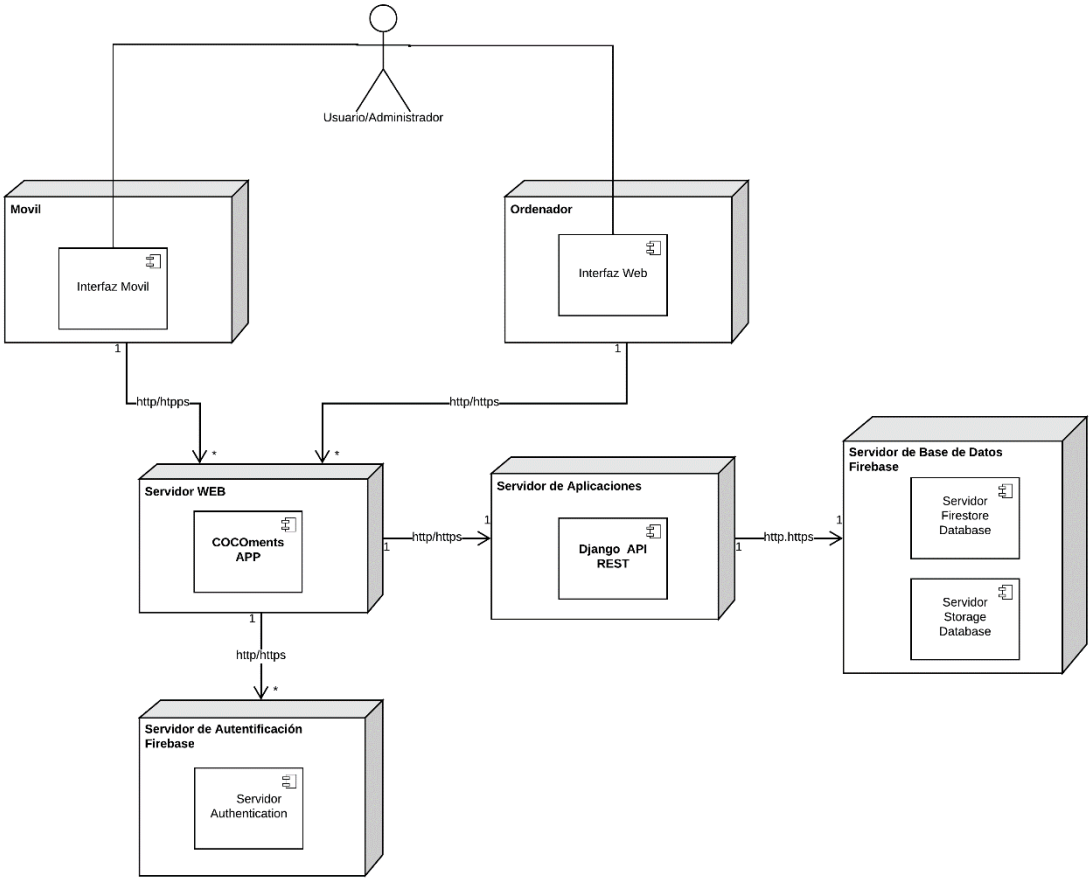


Figura 15. Diagrama de Despliegue de la aplicación hibrida

6.3.Diseño Detallado

En este apartado se presentan gráficos que ilustran cómo funciona el proceso interno de la aplicación para los distintos módulos desarrollados.

6.3.1. Diagrama de Actividades

En esta sección se presenta un total de seis diagramas de actividades, el primero referente a inicio de sesión, el segundo el registro de usuarios, la tercera creación de comentarios, cuarto listar comentarios para administrador, quinto reporte de comentarios del administrador y finalmente el sexto presenta el reporte de comentarios extraídos de redes sociales.

1. Inicio de Sesión

A continuación, se muestra el diagrama de actividades para el proceso de inicio de sesión de los usuarios.

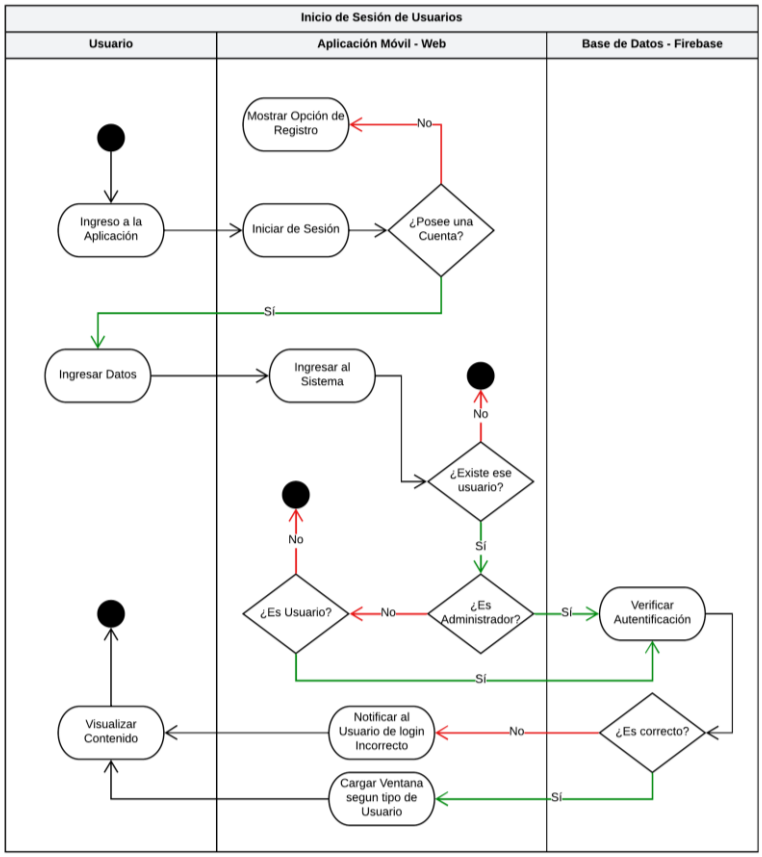


Figura 16. Diagrama de Actividades modulo Inicio de Sesión

2. Registro de Usuarios

A continuación, se muestra el diagrama de actividades para el proceso de registro de usuarios dentro de nuestra aplicación.

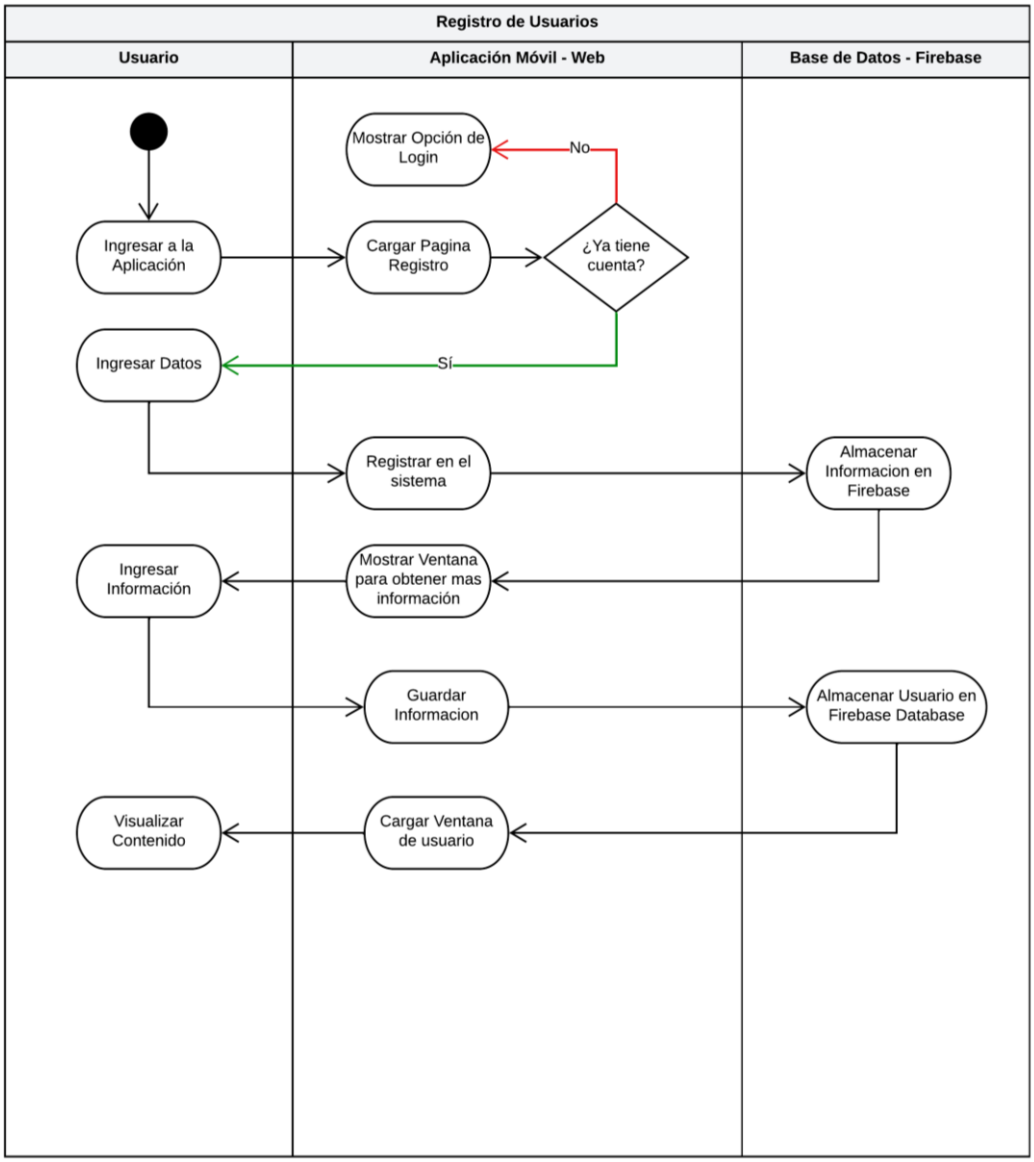


Figura 17. Diagrama de Actividades modulo Registro

3. Creación de Comentario

A continuación, se muestra el diagrama de actividades para el proceso de que un usuario pueda insertar un comentario en la aplicación.

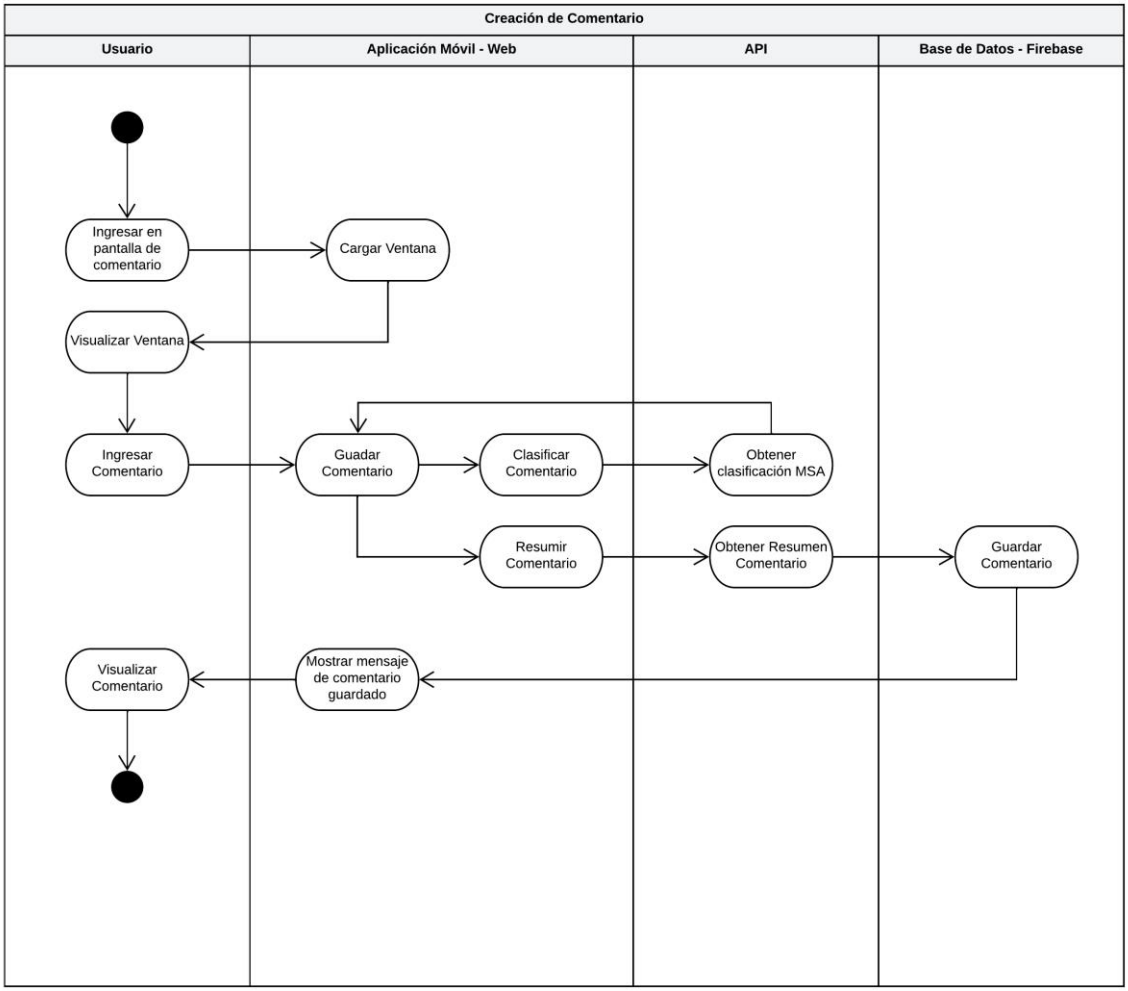


Figura 18. Diagrama de Actividades modulo creación de Comentario

4. Listar comentarios para Administrador

A continuación, se muestra el diagrama de actividades para el proceso de listar los comentarios que se encuentran almacenados dentro de Firebase.

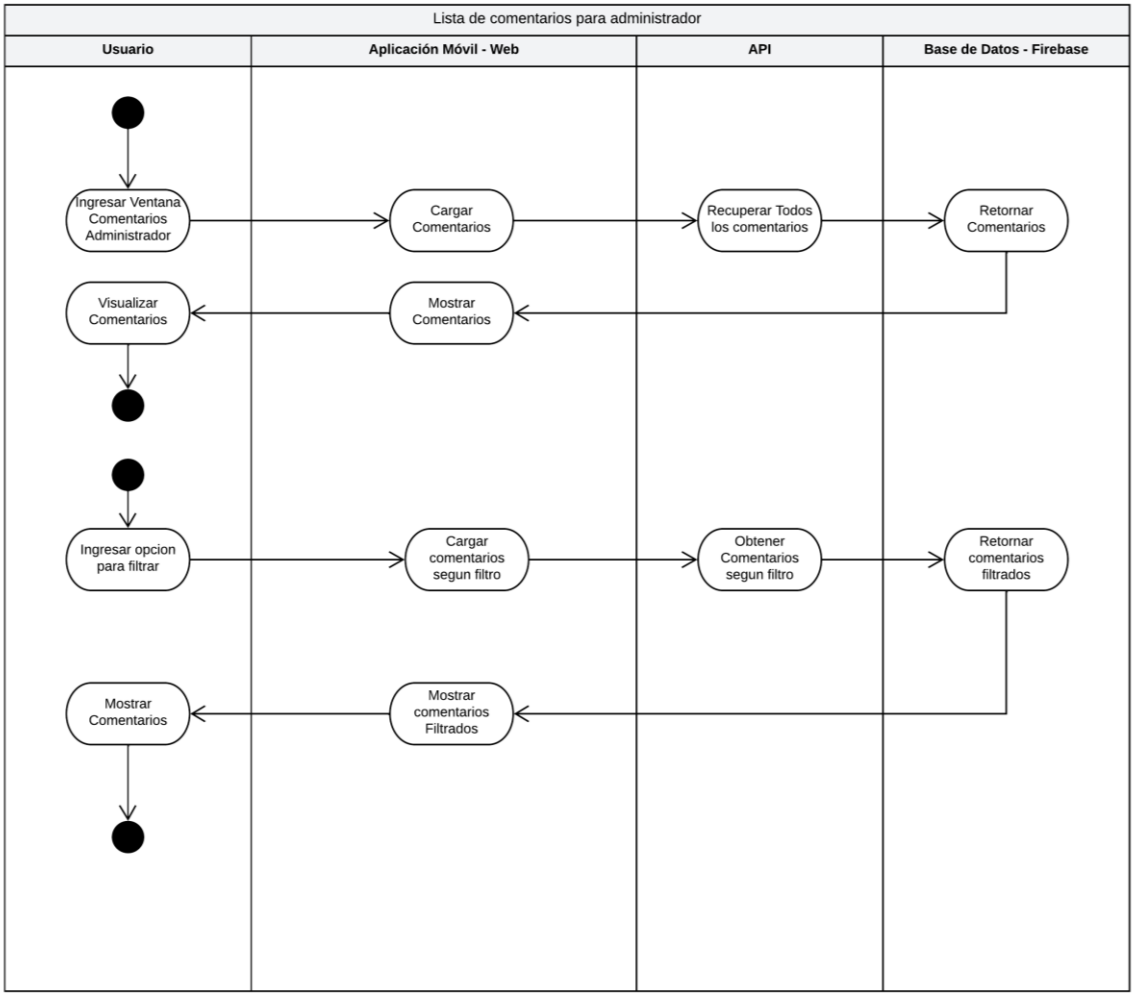


Figura 19. Diagrama de Actividades modulo listado de comentarios

5. Reporte comentarios Administrador

A continuación, se muestra el diagrama de actividades para el proceso de generación de reportes a través de gráficos, consultas de información de los comentarios que se encuentran almacenados dentro de Firebase.

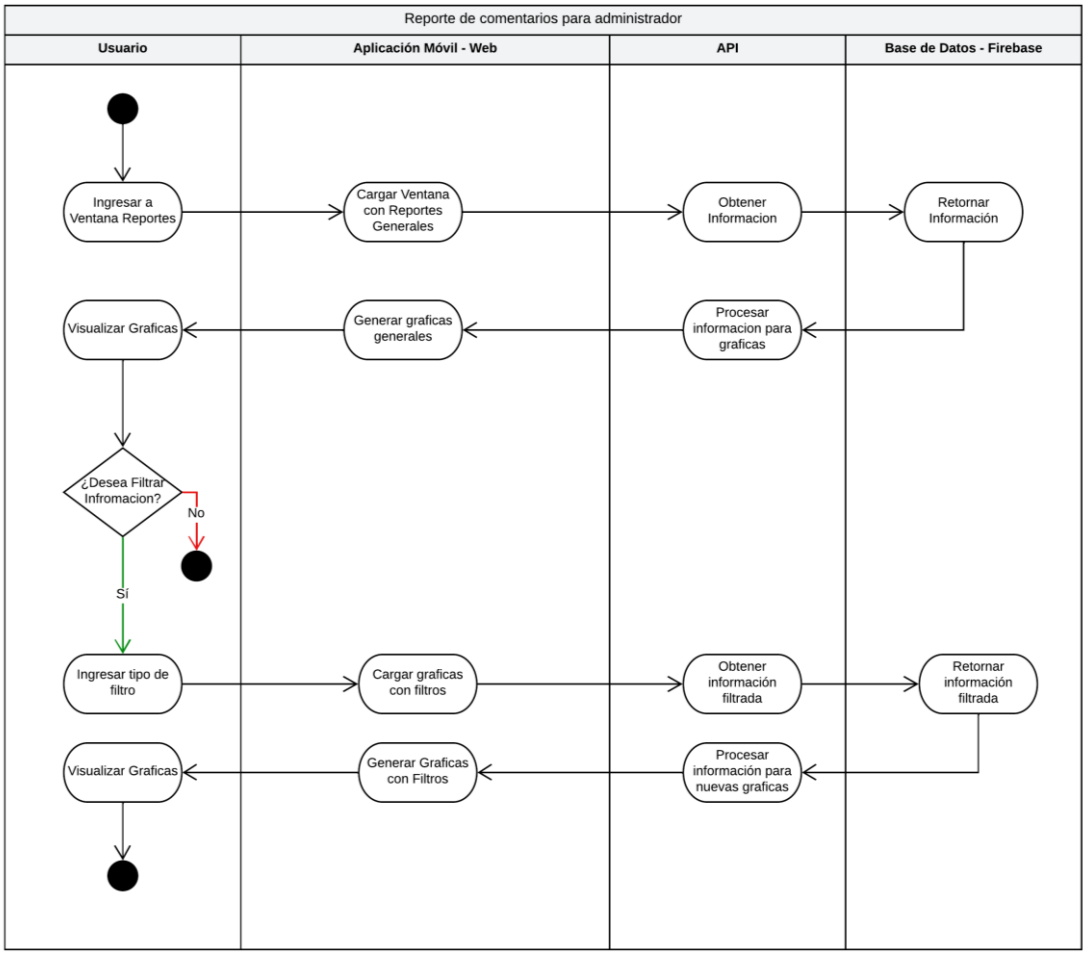


Figura 20. Diagrama de Actividades modulo Reportes de Comentarios

6. Reporte comentarios extraídos de redes sociales

A continuación, se muestra el diagrama de actividades para el proceso de generación de reportes de comentarios extraídos de redes sociales.

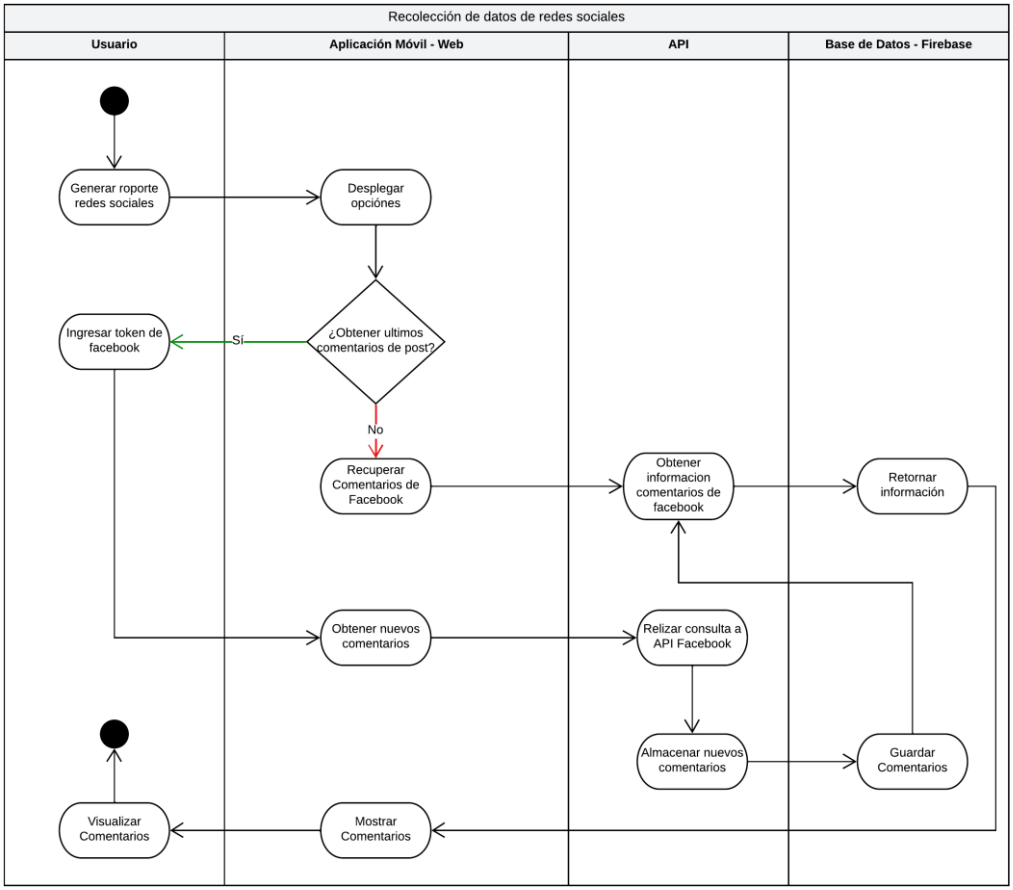


Figura 21. Diagrama de Actividades módulo de reportes de comentarios extraído desde redes sociales

6.3.2. Diagrama de Clases

En esta sección se presenta el diagrama de clases a fin de mostrar la estructura existente dentro de la aplicación, como las relaciones que existen entre las entidades.

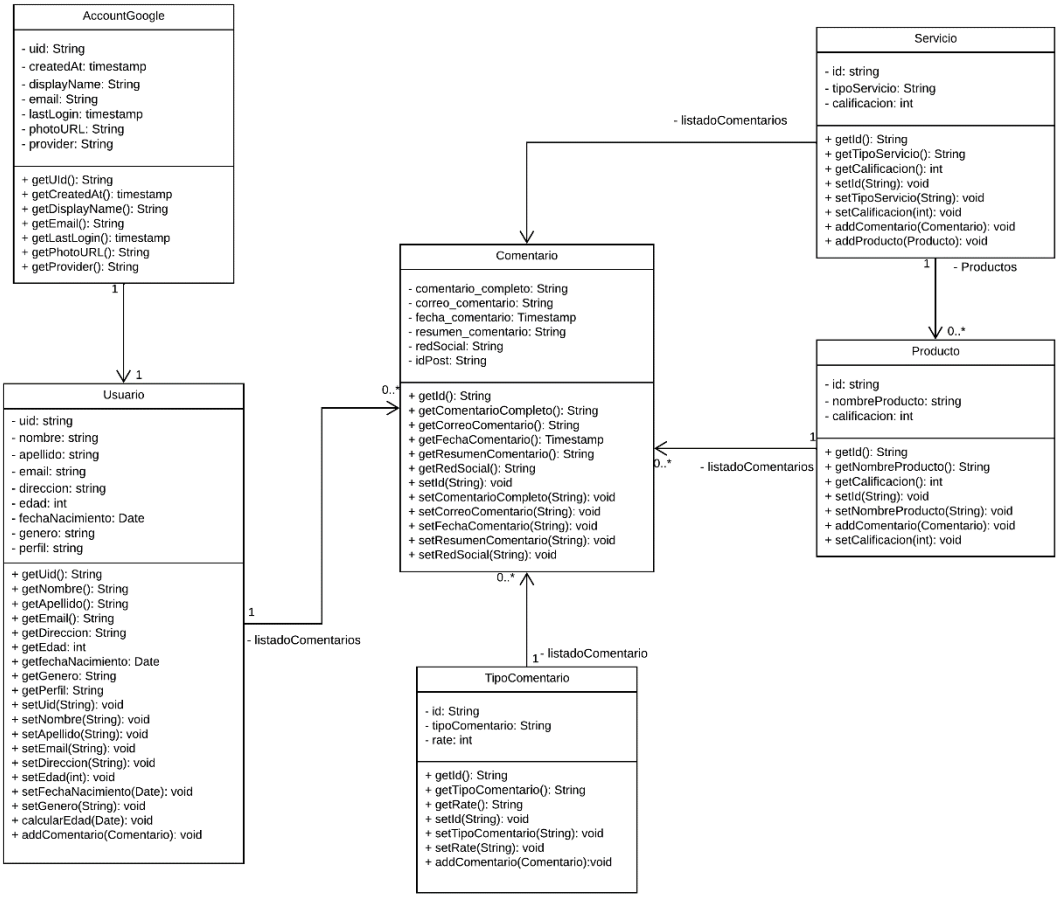


Figura 22. Diagrama de Clase de la Aplicación enfocada a Comentarios.

6.4.Desarrollo del Modelo de Análisis

En esta sección se presenta modelo de Análisis que tiene un enfoque en 3 fases, como se muestra en la siguiente.

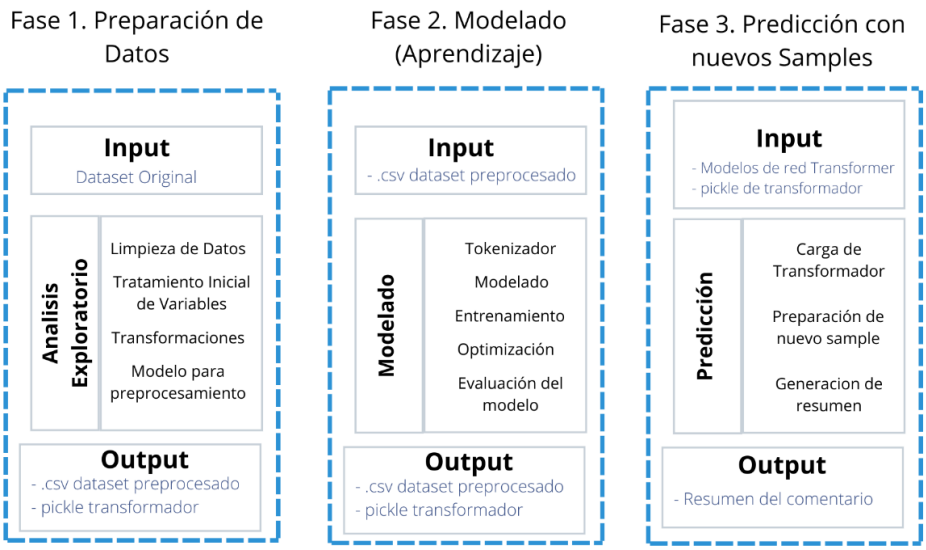


Figura 23. Modelo en fases para construcción de Transformer enfocado en el resumen de texto

6.4.1. Modelado

En la siguiente sección se abordará cada fase anteriormente presentada, dando a conocer los resultados obtenidos en cada fase y su proceso. Se presenta de la siguiente forma, la primera fase representa la preparación de datos, segunda fase el modelado de la Red Neuronal Transformers y finalmente la fase de predicción de nuevos samples.

A. Fase 1. Preparación de Datos

Para este aportado se tiene una entrada de conjunto de datos con 51716 reseñas acerca de restaurantes, dentro de este conjunto contiene un comentario acerca de la experiencia del cliente en el restaurante, este puede estar realizado con la comida, restaurante, bar, ambientación, entre otros; y un valor de sentimiento en donde cero se toma como que no le gusto y uno que si le gusto el establecimiento. Sin embargo, como se observa, no se tiene un resumen, por lo que esta parte fue realizada por nosotros, tomando 200 registros del dataset, 100 positivos y 100 negativos; para cada uno de ellos se generó su respectivo resumen, se generó un resumen más extenso y uno más corto, teniendo ahora tres columnas, comentario, sentimiento y resumen largo y resumen corto.

Posterior se procedió a realizar una limpieza del dataset eliminando valores nulos, eliminación de caracteres especiales, quitar espacios y tabulaciones, eliminación de palabras cortas, entre otras; este nuevo dataset limpio será con el que se trabajara dentro del modelo, sin embargo, para realizar un análisis de exploratorio se hace uso de técnicas de procesamiento de lenguaje natural.

Como primer punto se realiza una eliminación de stopwords, con el fin de eliminar palabras que siempre aparecen dentro del texto. A continuación, se muestran las palabras con más frecuencia dentro del dataset.

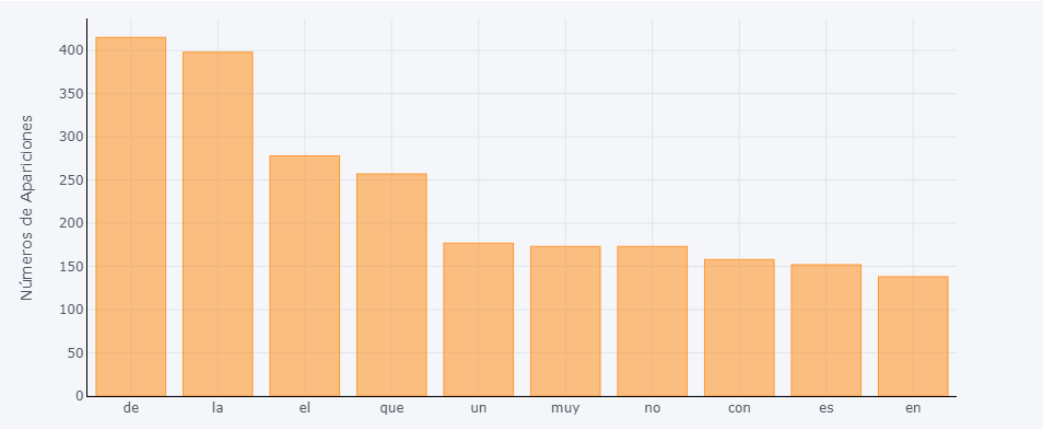


Figura 24. Palabras con más apariciones dentro del dataset

Después de realizar un procesamiento de eliminación de stopwords estas son las palabras ahora que más aparecen dentro del dataset.

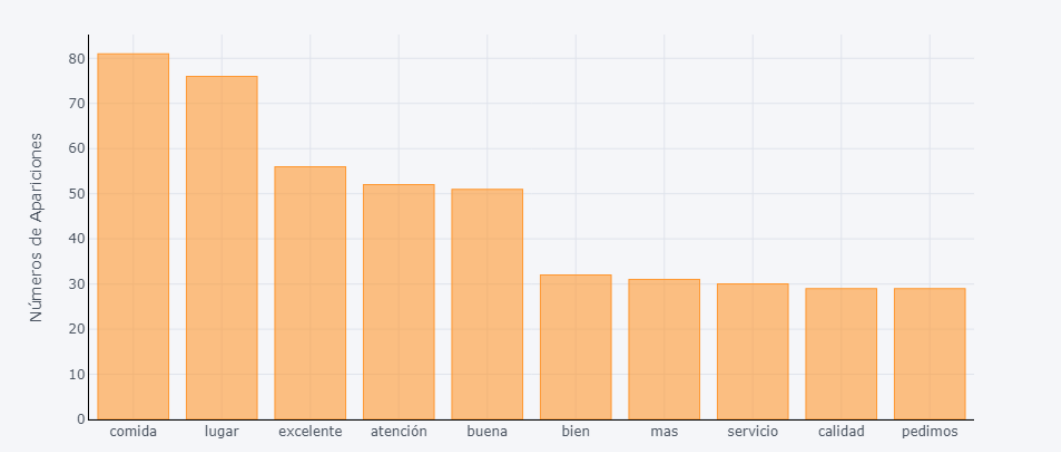


Figura 25. Palabras con más apariciones dentro del dataset una vez realizado un procesamiento de stopwords

Como se puede observar la diferencia es enorme, pasamos de tener palabras comunes a algo mucho más interesante para el negocio, en este caso al de bares y restaurantes. Pero esto es mucho más interesante cuando llevamos nuestro modelo a observar ngramas con más apariciones dentro del dataset. A continuación, se muestran los bigramas dentro del comentario sin haber aplicado stopwords.

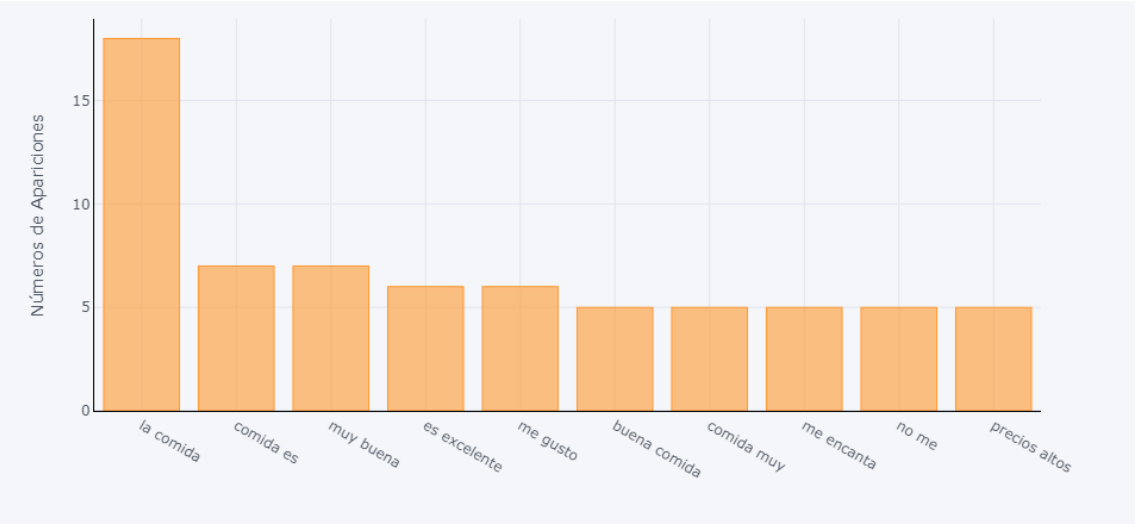


Figura 26. Bigrama del Dataset sin aplicar stopwords

Como se observa, al no aplicar eliminación de stopwords los bigramas no tienen mucho sentido y tampoco nos dice mucho, sin embargo, cuando se aplica eliminación de stopwords, se puede apreciar de mejor manera a lo que un comentario hace referencia con mayor frecuencia.

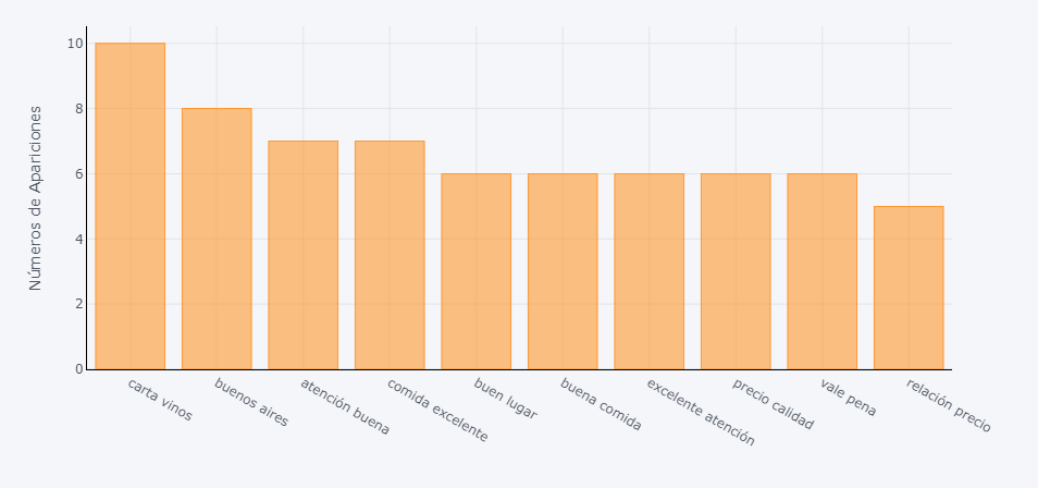


Figura 27. Bigrama del Dataset aplicando eliminación de stopwords

Dentro del Archivo también se analiza la frecuencia de las palabras una vez eliminado las stopwords, con esto se ha generado de manera visual una nube de palabras para

mayor entendimiento acerca de las palabras que más se encuentran dentro del dataset, ver imagen dentro de la sección de Anexos.

Uno de los parámetros más importantes es saber la longitud máxima que vamos a definir para el comentario y resumen, para esto se toma como referencia la longitud de los comentarios y resúmenes y se estableció que la media estaría entre 150-200 y 10 respectivamente, siempre y cuando se trabaje con los resúmenes cortos. Por último, se genera un modelo en donde se aplica técnicas de procesamiento de lenguaje natural y limpieza de texto; además, también se tiene un archivo que servirá de dataset para la próxima fase.

B. Fase 2. Modelado

En la fase dos, se utiliza Google Colab con GPU, esto debido a que brinda mayor rapidez de procesamiento y tarjetas gráficas a disposición para entrenar modelos, en la sección de Anexos se encuentra partes importantes del código fuente generado.

Como primer punto se tiene la división del dataset, el 90 por ciento destinado a entrenamiento, siendo 180 registros y 10 por ciento para evaluación del modelo, con 20 registros. También se define las longitudes máximas de caracteres para comentario y resumen 200 y 10 respectivamente.

El algoritmo está escrito en lenguaje de programación python, a través de pytorch, en donde se definió un primer modelo para realizar la tokenización de palabras tanto de comentarios como de resúmenes.

Para esto se utiliza el modelo ya entrenado de T5 (T5Tokenizer), se define una longitud máxima para el tokenizador tanto para comentarios como resumen, 1024 y 512 respectivamente. Este modelo nos retorna un diccionario con vectores de nuestras palabras ya tokenizadas, también una máscara de atención, un vector de etiquetas y una máscara de etiquetas.

A continuación, se crea otro modelo el cual recibe los datos de entrenamiento y evaluación y demás parámetros para su respectiva configuración, este modelo será el encargado de llamar al modelo de tokenizacion y enviar los conjuntos de datos; posterior el modelo procede a retornar información a través de dataloaders, que son vectores de importación de datos a gran escala, dentro de estos estarán cargados nuestros vectores de tokenización.

Ahora se crea un modelo para el transformador, este recibe el modelo, el tipo de tokenizador, y otros parámetros para almacenamiento y comportamiento. Aquí se obtiene parámetros de configuración de nuestro modelo, entradas, mascará de atención, etc. También en este punto se procede a tomar medidas de las salidas de entrenamiento y evaluación del modelo, también se procede a configurar el tipo de optimizador que tiene el modelo, en este caso se trabaja con Adam. Dentro de este modelo se lleva a cabo el entrenamiento y validación, los resultados se almacenan dentro de una carpeta outputs, la cual contendrá al modelo y su evolución durante

cada epoch generado, además de su valor de entrenamiento y validación. No se es posible almacenar todos los epochs generados debido al espacio que consume cada uno de estos, por esta razón solo se almacena el último epoch generado y este contendrá nuestro modelo con sus pesos y configuración.

Por último, se crea una clase la cual será la encargada de llamar a los distritos modelos anteriormente generados, esta clase será la encargada de recibir cada parámetro que se tiene para el modelo, datos de entrenamiento y evaluación, tamaño del tokenizador, tamaño del batch, número máximo de epoch, lugar de almacenamiento de nuestro modelo, entre otros. Esta clase es la encargada de mandar los datos de entrenamiento y validación a que sean tokenizados y posterior serán entrenados dentro del modelo T5. Contiene parámetros si se desea realizar parada temprana a nuestro modelo, si se desea trabajar con cpu o gpu. Una vez que se tiene todos estos parámetros configurados y listos, se procede a enviar a reentrenar nuestro modelo.

Con estos módulos o clases de pytorch, podemos reentrenar el modelo, para esto se utilizan los siguientes parámetros.

Tabla 14. Tabla parámetros para entrenamiento de modelo

Parametro	Descripción	Valor
train_df	Dataset para entrenamiento	180 registros
eval_df	Dataset para test y validación	20 registros
source_max_token_len	Longitud máxima para tokenizador de comentario	200
tarjet_max_token_len	Longitud máxima para tokenizador de resumen	18
batch_size	Número de muestras que se propagan a través de la red	12
max_epochs	Número de ciclos de entrenamiento	8
use_gpu	Habilitación de gpu para entrenamiento	True

Con estos parámetros de configuración se tienen los siguientes resultados.

Como primer punto, se tiene el comportamiento del modelo al momento de ser entrenado. Para esto se analiza las gráficas de perdida de entrenamiento (train loss) y perdida de validación (val_loss), se observa como el modelo de entrenamiento se va a justando a los datos, en el conjunto de validación se comienza con un valor más pequeño, pero con el paso del tiempo estos se llegan a igualas, con un rango menor

de diferencia, esto es debido a que el modelo ya está aprendiendo del conjunto de entrenamiento. Se puede concluir que se obtuvo un buen ajuste, no el mejor, pero válido, una manipulación de los parámetros y una mayor capacidad para procesamiento y un conjunto de datos de mayor tamaño permitirá mejorar considerablemente al modelo.

Graficas del comportamiento del modelo en las distintos epochs generados.

Tabla 15. Graficas de comportamiento a través de epochs y steps

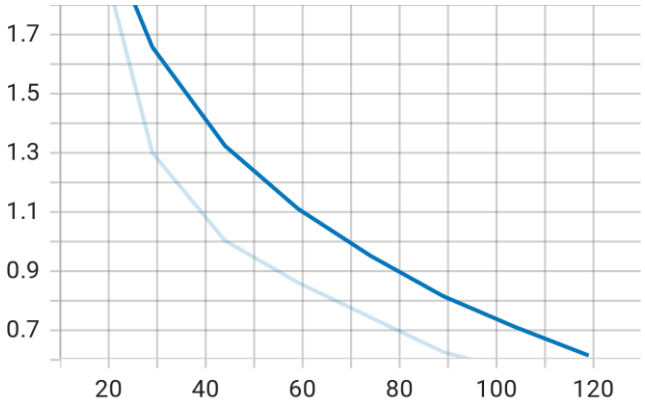


Figura 28. Grafica de perdida de entrenamiento (train los vs epochs)

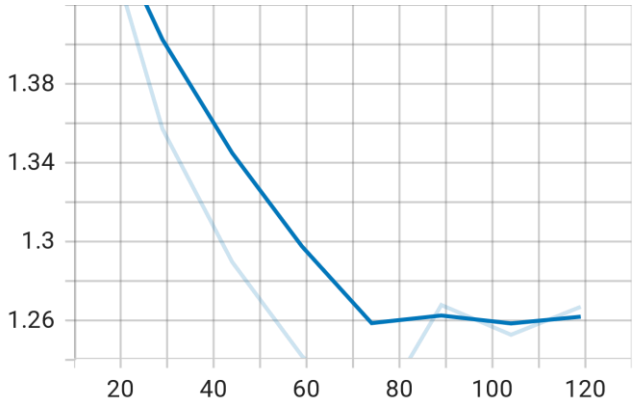


Figura 29. Grafica de perdida de validación (val los vs epochs)

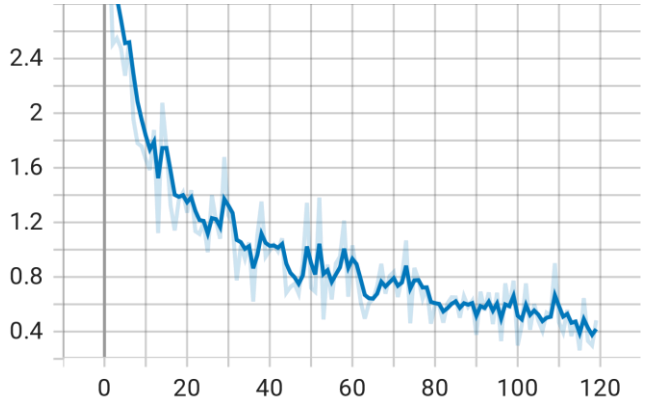


Figura 30. Grafica de perdida de entrenamiento (train los vs step)

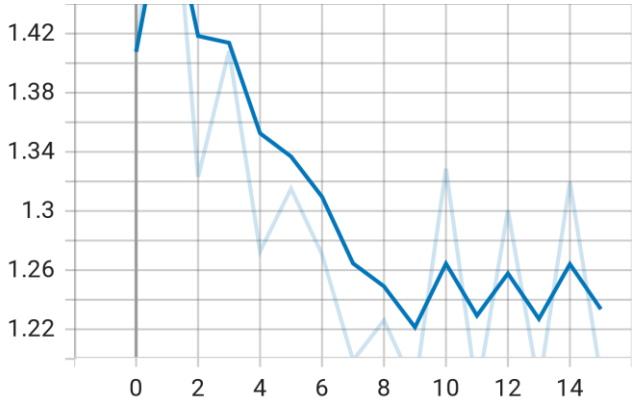


Figura 31. Grafica de perdida de validación (val los vs steps)

En segundo lugar, se analiza que tan buenos han sido los resúmenes generados por el modelo con respecto al conjunto de datos de evaluación, para esta tarea se hace uso de la métrica de evaluación ROUGE, la cual contiene múltiples métricas dentro de un solo lugar, en el modelo se evalúa Rouge 1, Rouge 2 y Rouge L.

Esta métrica mide el número de n-gramas coincidentes entre nuestro texto generado mediante el modelo y una referencia (resumen original). Por lo tanto, dentro de Rouge 1 se mide la tasa de coincidencia de unigramas entre la salida del modelo y el resumen original, esto es equivalente a Rouge 2, trabajando con bigramas. Por otro lado, Rogue L brinda la posibilidad de medir la subsecuencia más larga entre la salida del modelo y su referencia. Dentro de estas medidas se analizan precisión, recall y F1.

Recall

Aquí se cuenta el número de n-gramas superpuesto que se encuentra tanto en la salida del modelo como en la referencia, luego se divide este número por el número total de n-gramas en la referencia.

$$recall = \frac{cantidad\ de\ ngramas\ iguales\ en\ resumen\ generado\ y\ original}{numero\ de\ ngramas\ en\ resumen\ original}$$

Esta medida permite revisar si nuestro modelo es capaz de capturar información de nuestro comentario, sin embargo, si este valor tiene a ser alto, nos indica que nuestro modelo solo está extrayendo información, con el fin de tener una mejor puntuación.

Precisión

Para solventar el inconveniente anterior se hace uso de la métrica de precisión, en donde en lugar de dividir para la cantidad de ngramas que tiene la referencia, lo hacemos para el resumen generado por nuestro modelo.

$$presicion = \frac{cantidad\ de\ ngramas\ iguales\ en\ resumen\ generado\ y\ original}{numero\ de\ ngramas\ en\ resumen\ generado}$$

Puntuación F1

Con estos dos valores es sencillo calcular el valor de F1, se utiliza la siguiente fórmula.

$$F1 = 2 \times \frac{presicion \times recall}{presicion + recall}$$

Este valor nos va a servir de referencia, ya que este nos asegura que nuestro modelo no solo se base en capturar palabras, sino que lo haga sin llegar a generar palabras irrelevantes dentro del resumen esperado.

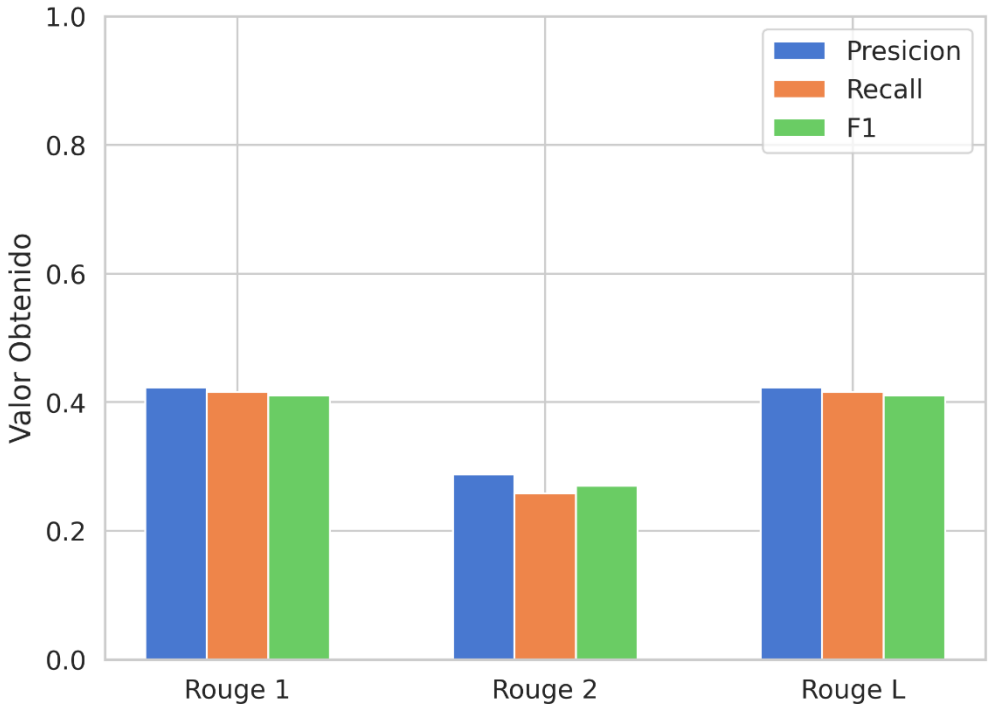


Figura 32. Rouge-N del Modelo generado para resumen de texto

Se observa en la anterior imagen nuestro modelo, con sus respectivas métricas de evaluación, en donde para Rouge – 1 se tiene un puntaje para F1 de 0.41; mientras que para Rouge – 2 se registró un 0.27 y por último para Rouge – L se obtuvo un valor de 0.41, pudiendo estos ser mejorados, con un mayor conjunto de datos y mejores resúmenes para su aprendizaje, así como también mejorar en el apartado de entrenamiento con más epochs para analizar si el modelo llega a aprender de mejor manera. Los puntajes obtenidos están correctos para nuestro propósito. Pueden llegar

a parecer pequeños, sin embargo, con este puntaje se llegan a obtener muy buenos resultados, a continuación, se muestra 4 ejemplos de comentarios generados, si desea ver todo el conjunto lo puede encontrar dentro de anexos.

Comentario	Resumen original	Resumen generado
la comida es excelente lo unico un poco restringida la opción para celiacos que podemos comer ensaladas muy variadas y solo algunas entradas como la polenta grillada con hongos que es deliciosa la atención es inmejorable y el ambiente es distinto y moderno	comida es excelente	comida es excelente
fuimos con mi novia el salón era extremadamente caluroso no prendieron al aire acondicionado demoraron aprox hora en traer los platos mi novia pidio un wok que consistia en solo arroz blanco y seco con verduras yo pedí un chivito que aprobaba raspando la atención es regular para abajo y la comida olvidable no lo recomiendo	experiencia horrible	atención es regular para abajo
para redondear todo es de mal gusto salvo la camarera que nos atendio bien el resto desastre precios altisimos la comida no es mas de lo que se puede comer en un bodegón italiano recomiendo que no vayan van a salir de mal humor	precios altisimos	precios altisimos
excelente restaurante fuimos personas pedimos dos platos para compartir una copa de vino bebidas cafe y como era un cumpleaños nos regalaron un increible tiramisú de postre el risotto es el plato mas rico que hay salió todo a la perfección hermoso lugar excelente atención y la iluminación es magnifica volveré	excelente restaurante	excelente restaurante

Figura 33. Resúmenes generados por nuestro modelo aplicando Transformer T5

Por último, esta fase nos da el modelo ya entrenado con sus pesos, tokenizador y estructura, estos archivos están almacenados en la carpeta de salida(output) del modelo, este directorio será utilizando en la siguiente fase, con el fin de poder realizar predicciones y obtener resúmenes para nuestros comentarios.

C. Predicción con nuevos samples

En esta sección el objetivo es llevar nuestro modelo anteriormente generado, a poder realizar predicciones para nuevas entradas de comentarios y ver como este se comporta ante este nuevo escenario.

El modelo fue entrenado dentro de Google Colab por lo tanto, ahora tenemos que configurar e instalar los paquetes necesarios para que el modelo, pueda llegar a correr dentro de una computadora propia, en este punto se trabaja sin GPU, sin embargo dentro de la codificación del algoritmo se ha habilitado una opción para la predicción de comentarios con el uso de tarjetas gráficas, esto con el fin de acelerar el proceso. En nuestro caso vamos a trabajar con CPU para este punto.

En esta sección se tiene como entrada al modelo generado y el proceso de transformación de texto, estos dos archivos son necesarios para el proceso. Se crea una clase en donde se define nuestra función de predicción y carga de modelo. Estas dos funciones van de la mano, primero se carga el modelo, una vez que se tiene el modelo y tokenizador listo se procese a realizar la predicción, la cual tiene los siguientes parámetros de configuraciones.

Tabla 16. Parámetros para predicción de nuevos comentarios con modelo T5

Parámetro	Descripción	Valor
source_text	Comentario ingresado	Cadena de texto
max_length	Longitud máxima para el comentario	32 caracteres
num_return_sequences	Numero de predicciones que se devuelven	1
num_beams	Valores por defecto	2
top_k		50
top_p		0.95
do_sample		True
repetition_penalty		2.5
length_penalty		1.0
early_stopping		True
skip_special_tokens		True
clean_up_tokenization_spaces		True

Este método nos retorna nuestro comentario ya generado, para ver los resultados a continuación se muestra 4 ejemplos que fueron realizados utilizando nuestro modelo, los comentarios han sido tomados de la página de Tripadvisor que hacen referencia a comentarios sobre restaurantes.

En la siguiente imagen se observa como nuestro modelo es capaz de generar un resumen pequeño, debido a que nuestro dataset contenía comentarios supercortos y fueron entrenados a partir de esta información, sin embargo, el resumen que se obtiene es bastante bueno, está enfocado en el contexto detrás del comentario ingresado por el

usuario, como se demostró en las medidas de calidad, en algunos casos llega a fallar o no a entender el contexto. Después de realizar algunos ejemplos, también se detecta que el algoritmo no trabaja bien con comentarios cortos, en este aspecto lo que hará es copiar literal el texto.

Comentario	Resumen Generado
Un lugar muy bonito, con mucha decoración, agradable ambiente. La atención es muy demorada y la comida no es lo que se espera.	atención agradable ambiente
La comida deliciosa! El ambiente hermoso, la atención excelente, y todo por precios razonables! Definitivamente recomendado!	ambiente hermoso
El ambiente de lujo, la atención esmerada y con mucha amabilidad, la comida con una combinación de sencillez y que resalta cada sabor en la boca. No soy de comer postres, pero me dieron uno de crema de menta que es un real éxtasis.	ambiente de lujo
El restaurante El Jardin del Hotel Victoria, es un lugar que conserva una tradición de más de 40 años en Cuenca, pues este fue el primer restaurante de servicio completo que abrió en la ciudad y que se mantiene hasta la actualidad. Una carta bien lograda, con platos esenciales y bien servidos siguiendo principios culinarios de siempre. El servicio es muy bueno, y basta haber ido más d una vez que la relación con sus dueños y empleados se vuelve absolutamente familiar. La vista del Tomebamba y del jardín del Hotel representan un valor agregado enorme para las personas que buscan un remanso pacífico en sus labores. Muy recomendables las sugerencias diarias del chef y las propuestas gastronómicas de temporada. Actualmente se encuentra solo con servicio a domicilio, esperado que vuelva pronto a abrir sus puertas según comentan sus propietarios.	restaurante servicio bueno
Llegamos con muy altas expectativas.A las 13h45 ingresamos, nos tomaron el pedido a las 14h00, las entradas las trajeron a las 14h40, luego de eso los platos fuertes llegaron a la 15h25, y cuando llegó la comida fue muy decepcionante, platos desabridos que ni siquiera entran dentro de un gusto normal... finalmente las bebidas nunca llegaron. Uno de los meseros a mi parecer dio un servicio quemimportista.Sentí que boté el dinero a la basura.No volveré a este lugar, lástima que la vista es bonita.	comida fue muy decepcionante

Figura 34. Resúmenes generados por el Modelo T5

6.4.2. Implementación y despliegue como servicio

En esta sección, se integra nuestro modelo dentro de la aplicación híbrida inteligente, para esto se generó dos archivos de configuración, uno enfocado al modelo y otro para el procesamiento y limpieza de comentarios, además se incluye directorio con el modelo entrenado. Con estos dos archivos se crea un servicio rest, con Django para poder consumir nuestro modelo. A continuación, se muestra al servicio ya configurado y realizando pruebas con el software postman. Para esto nuestro servicio recibe tres parámetros: comentario, correo de la persona que realiza el comentario y el tipo de comentario.

```
1   var request = require('request');
2   var options = {
3     'method': 'POST',
4     'url': 'http://192.168.100.184:8080/api/guardarComentario',
5     'headers': {
6       'Content-Type': 'application/json'
7     },
8     body: JSON.stringify({
9       "contenido_comentario": " El resto de la comida en la misma línea, restaurante
            caro y regular. Creo que como punto de encuentro para tomar una cerveza
            interesante",
10      "tipoComentario": "mixed",
11      "correoComentario": "barrerajuan930@gmail.com"
12    })
13
14  };
15  request(options, function (error, response) {
16    if (error) throw new Error(error);
17    console.log(response.body);
18  });
19
```

Figura 35. Parámetros servicio rest para consumo de modelo Transformer

El almacenamiento, se realiza en Django, almacenando el comentario con su respectiva información dentro de firebase.

6.5. Codificación

En esta sección se explicará cómo se realiza la implementación de nuestra aplicación, tanto la capa de usuario, negocio y de datos y el modelo de análisis de datos propuesto. Se estructura de la siguiente manera, primero se presenta el modelo de análisis, segundo la aplicación híbrida referente a la capa de usuario, tercero la aplicación híbrida referente a la capa de negocio, finalmente se tiene la codificación de Web Services.

6.5.1. Modelo Análisis

Para el desarrollo de este modelo, se utiliza el lenguaje de programación Python, haciendo uso de cuadernos de Jupyter notebook proporcionado a través del conjunto de librerías de Anaconda, para el entrenamiento del modelo, se hace uso de Google Colab.

A continuación, se mostrará la estructura básica del modelo implementado. Se observan 3 archivos en donde estará la información de cómo se codificó el modelo, la fase 1 es el archivo de análisis del dataset, en donde encuentra la limpieza, tratamiento de variables, transformador con funciones de limpieza de texto y preprocesamiento de lenguaje natural; el archivo de la fase 2, en cambio, contendrá al tokenización del dataset, modelado, entrenamiento, optimización y evaluación del modelo; y por último el archivo para la fase 3 que contiene la carga de transformador de texto que realiza la limpieza, carga del modelo pre-entrenado en la fase 2, función para predicción y generación de resúmenes de muestra de la eficacia del modelo.

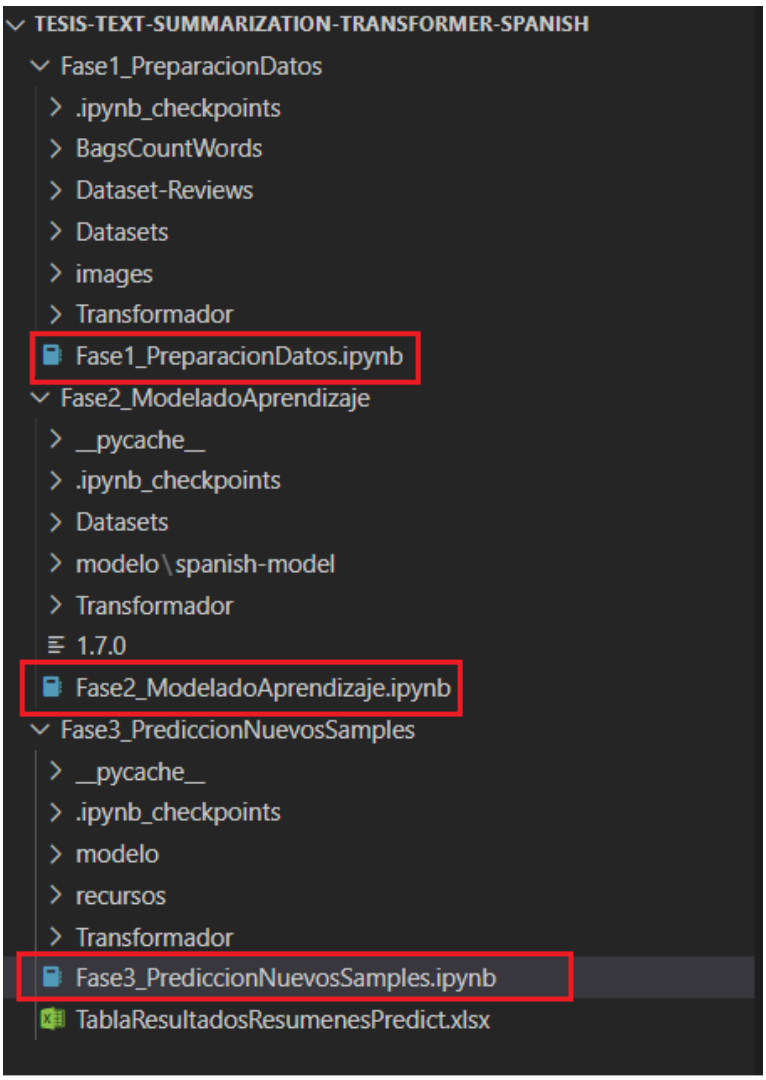


Figura 36. Estructura organizacional de archivos del algoritmo

En cada una de estas fases se almacenan los resultados que se han obtenido durante su proceso, estos se almacenan en cada directorio de cada fase. Estos archivos se encuentran como código abierto (open source), dentro de nuestro GitHub, el cual permite el trabajo colaborativo dentro del desarrollo del proyecto, además de su control de versiones. A continuación, se adjunta el enlace para acceder a los recursos del algoritmo. Si quiere analizar el código puede acceder a la carpeta de modelo de la red transformer (Modelo-Text-Summarization-Transformer-Spanish) dentro de este directorio tendrá la posibilidad de revisar como fueron las fases del algoritmo y los resultados se obtenidos en cada una.

Código fuerte:

- https://github.com/Juancarlos56/TESIS-ALGORITMO-TRANSFORMER-APLICACION_HIBRIDA_PARA_RESUMIR_Y_REPRESENTAR_COMENTARIOS.git

6.5.2. Aplicación Híbrida FrondEnd

En esta sección se presenta la codificación para la aplicación híbrida, para el desarrollo de la capa de usuario, se utiliza el lenguaje de programación Typescript a través de IONIC. Como herramienta de desarrollo se usa Visual Studio Code. A continuación, se muestra la estructura de la aplicación, antes de ser desplegada.

Es de mencionar que la aplicación fue desarrollada con IONIC 6, y angular 13. A continuación se muestra los requerimientos que se necesita para la aplicación. Previamente, tener instalado Node.js LTS v16.14.2.

```
Ionic:

   Ionic CLI                     : 6.19.0
   Ionic Framework               : @ionic/angular 6.1.3
   @angular-devkit/build-angular : 13.2.6
   @angular-devkit/schematics    : 13.2.6
   @angular/cli                  : 13.2.6
   @ionic/angular-toolkit        : 6.1.0

Capacitor:

   Capacitor CLI      : 3.5.0
   @capacitor/android : 3.5.0
   @capacitor/core    : 3.5.0
   @capacitor/ios     : 3.5.0

Utility:

   cordova-res : not installed globally
   native-run  : 1.5.0

System:

   NodeJS : v16.14.2 (C:\Program Files\nodejs\node.exe)
   npm    : 8.5.0
   OS     : Windows 10
```

Figura 37. Requerimientos para desarrollo de FrontEnd con IONIC 6

Explicación de estructura del FrontEnd de la aplicación.

La carpeta en donde que contiene todas las librerías necesarias para el correcto funcionamiento de la aplicación se encuentra dentro de node_modules, para instalar todas las dependencias del proyecto colocarse dentro de la carpeta raíz del proyecto y ejecutar: npm install. Esta línea le proporciona todas las librerías que requiere el proyecto.

Dentro del proyecto se tienen directorios destinados al usuario y al administrador, dentro de ellos se tiene los módulos que estos pueden efectuar, usuario: agregar más información, página de bienvenida, ingreso de comentarios, listado de los comentarios realizados, perfil de usuario; Administrador: reportes de comentarios, listado de todos los comentarios, reportes del restaurante/bar, reportes de redes sociales. Además de las páginas de login y registro. Los archivos de angular.json y package.json contendrán información sobre las versiones de librerías que se están utilizando y la estructura de configuración del proyecto.

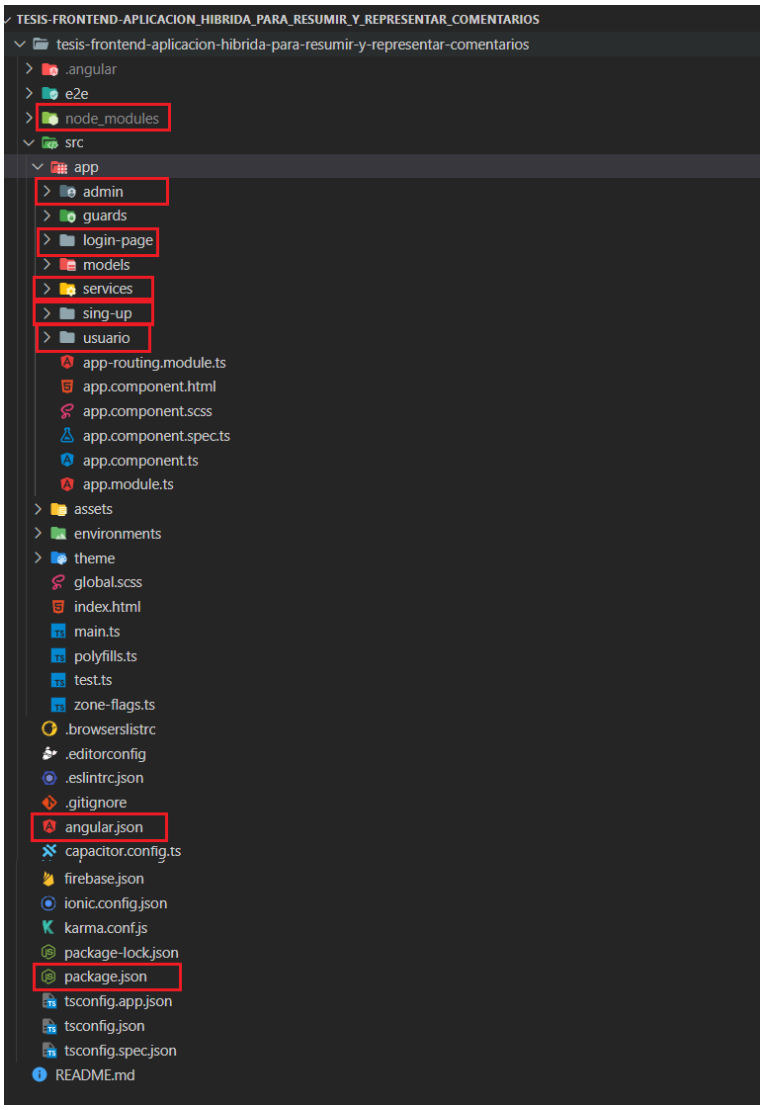


Figura 38. Estructura del FronEnd de la aplicación implementada sobre IONIC

A continuación, se adjunta el enlace para acceder a los recursos de la capa de usuario.

Código fuerte:

- https://github.com/Juancarlos56/TESIS-FRONTEND-APLICACION_HIBRIDA_PARA_RESUMIR_Y_REPRESENTAR_COMENTARIOS.git

6.5.3. Aplicación Hibrida BackEnd

En esta sección se presenta la codificación para la aplicación híbrida, para el desarrollo de la capa de negocio, se utiliza el lenguaje de programación Python a través del framework de Django. Como herramienta de desarrollo se usa Visual Studio Code. A continuación, se muestra la estructura de la aplicación, antes de ser desplegada.

Es de mencionar que la aplicación fue desarrollada con Django 4.0.4 y Django Rest Framework 3.13.1 además, de la librería pyrebase 3.0.27 y firebase-admin 5.2.0. Para esto se creó varios archivos que son necesarios para el correcto funcionamiento de la aplicación. Se tiene librerías dedicas a la aplicación, al algoritmo y a las redes sociales.

Explicación de estructura del BackEnd de la aplicación.

En esta sección se da una explicación detallada de la estructura que conforma el BackEnd manejado dentro de la aplicación, como también capturas de pantalla que muestran algunos de los servicios como serviceAccount.json al igual que una captura de la estructura creada.

Como primer punto se tiene al archivo de configuración de la aplicación (settings.py), dentro de este se configura la información del funcionamiento de la aplicación, los hosts permitidos, instalación de librerías y configuración de variables. También se agrega un archivo para recuperar el entorno de anaconda (env_Tesis.yaml) por si desea replicar el código, este archivo cuenta con todas las librerías utilizadas dentro de la configuración.

El directorio de claves privadas contendrá la información de Firabase para poder manipularla desde Python (serviceAccount.json), esto lo puedes obtener desde la consola de firebase. Los archivos .py son configuraciones para obtener variables globales de firebase y Facebook API.

El directorio Lógica, contiene el desarrollo del algoritmo para resumen de texto utilizando transformer T5 (ResourcesFilesTransformerSpanish), estos archivos buscarán a los pesos y configuración de la red dentro del directorio “static / modelo / TextSummarizationT5 /

nombre de modelo entrenado". Dentro de este directorio también se tiene la configuración para obtención y almacenamiento de comentarios utilizando redes sociales.

El directorio de Restful, contendrán todos los métodos desarrollados para el consumo de nuestra aplicación, ya sean consultas a la base de datos, registrar nuevos comentarios, obtener información del restaurante, obtener información de la red social proporcionada, entre otros.

El archivo de configuración de urls.py contendrá los paths y utls para la llamada desde el FrontEnd. El archivo enviroment.env contendrá la información de nuestra aplicación de Django y el token de Facebook de larga duración. Si desea instalar los paquetes se creó los archivos .txt para instalar todos los requerimientos necesarios.

Si quiere replicar el proyecto, tendría que crear dos archivos:

1. serviceAccount.json

   Agregar la siguiente configuración, información de su proyecto de firebase:

```
{
    "type": "",
    "project_id": "",
    "private_key_id": "",
    "private_key": "",
    "client_email": "",
    "client_id": "",
    "auth_uri": "",
    "token_uri": "",
    "auth_provider_x509_cert_url": "",
    "client_x509_cert_url": ""
}
```

Figura 39. Configuración para Firebase

2. enviroment.env

   Variables de configuración del proyecto de Django

```
SECRET_KEY='' # Del proyecto en Django
DEBUG=True
apiKey=""
messagingSenderId=""
appId=""
tokenAPI=""
```

Figura 40. Configuración para Django y Facebook API

A continuación, se adjunta el enlace para acceder a los recursos de la capa de negocio.

Código fuerte:

- https://github.com/Juancarlos56/TESIS-BACKEND-APLICACION_HIBRIDA_PARA_RESUMIR_Y_REPRESENTAR_COMENTARIOS.git

A continuación, se muestra la estructura del proyecto realizada en Django

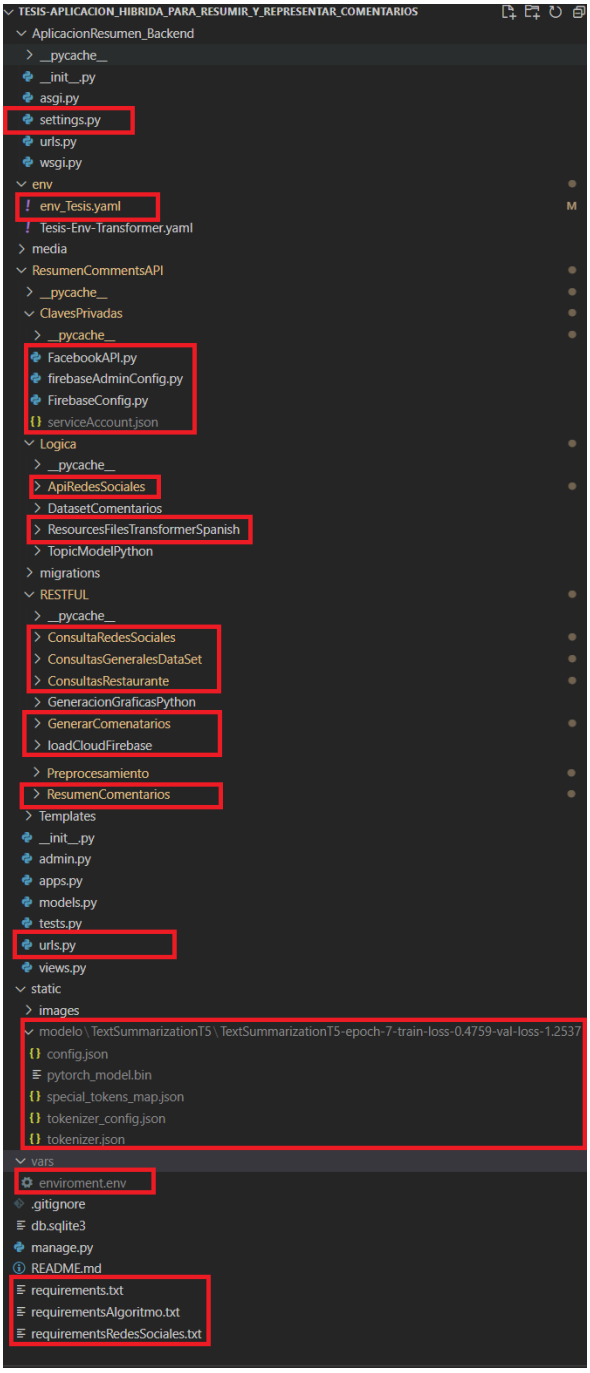


Figura 41. Estructura de Backend de la aplicación sobre Django

6.5.4. Codificación de Web Services

A continuación, observaremos el desarrollo del web service las cuales están en la parte del BackEnd en Django:

URL	Tipo	Explicación	Parámetros de Entrada/Salida
guardarComentario	POST	Nos permite almacenar nuestros comentarios dentro de firebase, internamente dentro del método se obtiene el resumen del comentario.	**Entrada**: comentario completo, tipo de comentario, correo de la persona que realiza el comentario. **Salida**: retorna si se almacenó o no un error.
topPalabrasComentario/ <str:tipo>	GET	Top de palabras con mayor frecuencia dependiendo del tipo de comentario dentro de la lista de comentarios	**Entrada**: tipo de comentario (very positive, positve,mixed, negative, very negative) **Salida**: path de la imagen generada y almacenada en storage
topPalabrasComentario/	GET	Top de palabras con mayor frecuencia dentro de la lista de comentarios	**Salida**: path de la imagen generada y almacenada en storage
topPalabrasComentario RangoFechas	POST	Top de palabras con mayor frecuencia dependiendo del rango de fecha proporcionado	**Entrada**: fecha de inicio, fecha de fin **Salida**: path de la imagen generada y almacenada en storage
topPalabrasResumen /<str:tipo>	GET	Top de palabras con mayor frecuencia dependiendo del tipo de comentario dentro de la lista de resúmenes	**Entrada**: tipo de comentario (very positive, positve,mixed, negative, very negative) **Salida**: path de la imagen generada y almacenada en storage

topPalabrasResumen	GET	Top de palabras con mayor frecuencia dentro de la lista de comentarios con respecto a los resúmenes	**Salida**: path de la imagen generada y almacenada en storage
topPalabrasResumen RangoFechas	POST	Top de palabras con mayor frecuencia dependiendo del rango de fecha proporcionado con respecto al resumen	**Entrada**: fecha de inicio, fecha de fin **Salida**: path de la imagen generada y almacenada en storage
graficasNgramas Palabras	POST	Nos permite analizar cuáles son los ngramas que más existen dentro de nuestro dataset de comentarios, tanto para resumen como texto completo.	**Entrada**: analizar comentario (completo, resumen), numero de palabras (cuantas palabras quiero obtener), número de ngramas (1,2,3), tipo de comentario. **Salida**: palabras y sus números de apariciones
generandoTopicDataset	POST	Analizar los tópicos más importantes dentro del dataset de comentarios, para resumen o texto completo.	**Entrada:** Tipo de comentario (resumen o texto completo) **Salida:** Lista de tópicos obtenidos con sus valores
buscarEnDataset	POST	Permite buscar palabras dentro del dataset se trabaja con texto completo o resumen	**Entrada:** tipo de comentario (texto completo o resumen), texto que desea buscar, clasificación del comentario (Opcional)
usersGenero	GET	Permite obtener cuantos hombres y mujeres tenemos dentro de nuestros usuarios registrados.	**Salida:** usuarios por sexo con su respectivo valor.
cantidadDe ComentariosPorTipo	GET	Permite obtener cuantos comentarios existen por clasificación de comentarios	**Salida:** lista de los tipos comentarios por clasificación con su respectivo valor.

usersEdad	GET	Obtener la edad de nuestros usuarios registrados dentro de la aplicación	**Salida:** lista por rango de edades de los usuarios registrados dentro de nuestra aplicación
comentariosUsuario	POST	Devuelve la lista de comentarios realizados por el usuario según su correo electrónico.	**Entrada:** Correo de usuario **Salida:** Lista de comentarios del usuario
listarComentarioFecha	POST	Lista de todos los comentarios filtrados por fecha.	**Entrada:** fecha de inicio y fecha de fin **Salida:** Lista de comentarios filtrada por fechas
listarComentario Sentimiento	POST	Lista de todos los comentarios filtrados por tipo de clasificación.	**Entrada:** tipo de clasificación **Salida:** Lista de comentarios filtrada por clasificación
listarComentario SentimientoFecha	POST	Lista de todos los comentarios filtrados por tipo de clasificación y fecha.	**Entrada:** tipo de clasificación, fecha inicio y fecha fin **Salida:** Lista de comentarios filtrados
cantidadDeComentariosPor TipoSentimientoEnTiempo	POST	Obtención de cuantos comentarios por fecha se realizaron	**Entrada:** fecha inicio y fecha fin **Salida:** Cantidad de comentarios por clasificación se obtienen en esa fecha
obtenerComentarios Facebook	GET	Obtener lista de últimas 25 publicaciones de nuestra página de Facebook	**Salida**: retorna la lista de comentarios de publicaciones

guardarComentarios Facebook	GET	Guardar lista de últimas 25 publicaciones de nuestra página de Facebook	**Salida**: retorna mensaje si se almacenó correctamente o uno un error.
filtrarComentarios FacebookFecha	POST	Devuelve la lista de comentarios dentro de las publicaciones de Facebook dentro de fechas específicas	**Entrada:** fecha inicio y fecha fin **Salida**: retorna la lista de comentarios de publicaciones en rango de fechas
obtenerMejores PlatosRestaunte	GET	Nos devuelve los mejores platos, y los más mencionados dentro de los comentarios con clasificación very positive y positive	**Salida**: Lista de platos con mejores calificaciones
obtenerPeores PlatosRestaunte	GET	Nos devuelve los peores platos, y los más mencionados dentro de los comentarios con clasificación very negative y negative	**Salida**: Lista de platos con peores calificaciones

Para observar el código de cada función puede revisar nuestro cosijo en GitHub, antes mencionado.

6.6. Pruebas Unitarias

En esta sección se presentan las pruebas unitarias creadas para mostrar el resultado de la aplicación, son basados en los requerimientos funcionales extraídos con anterioridad.

Tabla 17. Prueba Unitaria para inicio de sesión

Prueba funcional 1	Código	PF-01	Resultado
	Responsable	KB, RH	Aprobada
	Fecha	13/07/2022	

Requerimiento:

RF-01 Iniciar sesión de usuarios

Resultado Esperado:

El usuario ingresa correo y contraseña dentro de las cajas de texto, se validan los datos, en caso de que los datos sean correctos entonces la aplicación abre las páginas creadas para usuario.
En caso de ingresar mediante Google el usuario debe presionar sobre el botón con el icono de Google y posteriormente se abre las páginas autorizadas para el usuario cliente.

Resultado Obtenido:

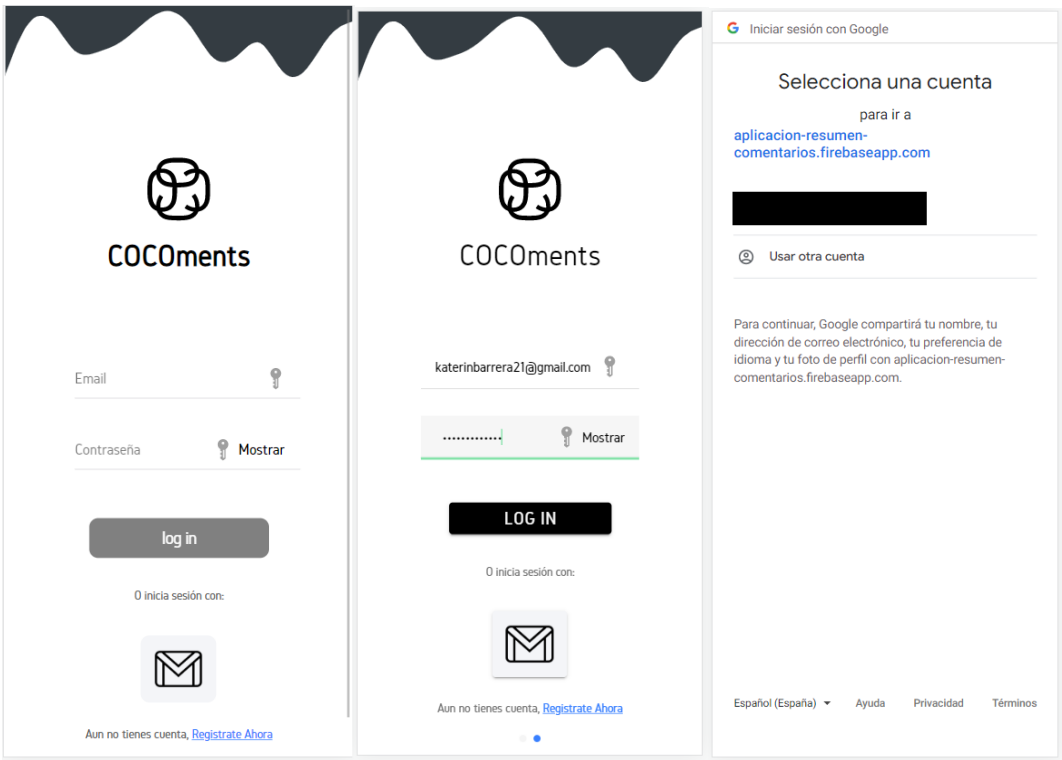


Tabla 18. Prueba Unitaria para Registro de usuario/cliente

Prueba funcional 2	Código	PF-0	Resultado
	Responsable	KB, RH	Aprobada
	Fecha	13/07/2022	

Requerimiento:

RF-01 Registrar usuario/cliente

Resultado Esperado:

El usuario cliente tiene la opción de registrarse dentro de la aplicación. Primero se muestra una página ingrese su correo y contraseña, posterior una página para que ingrese información adicional relevante para la aplicación. El usuario también puede registrarse mediante Google.

Resultado Obtenido:

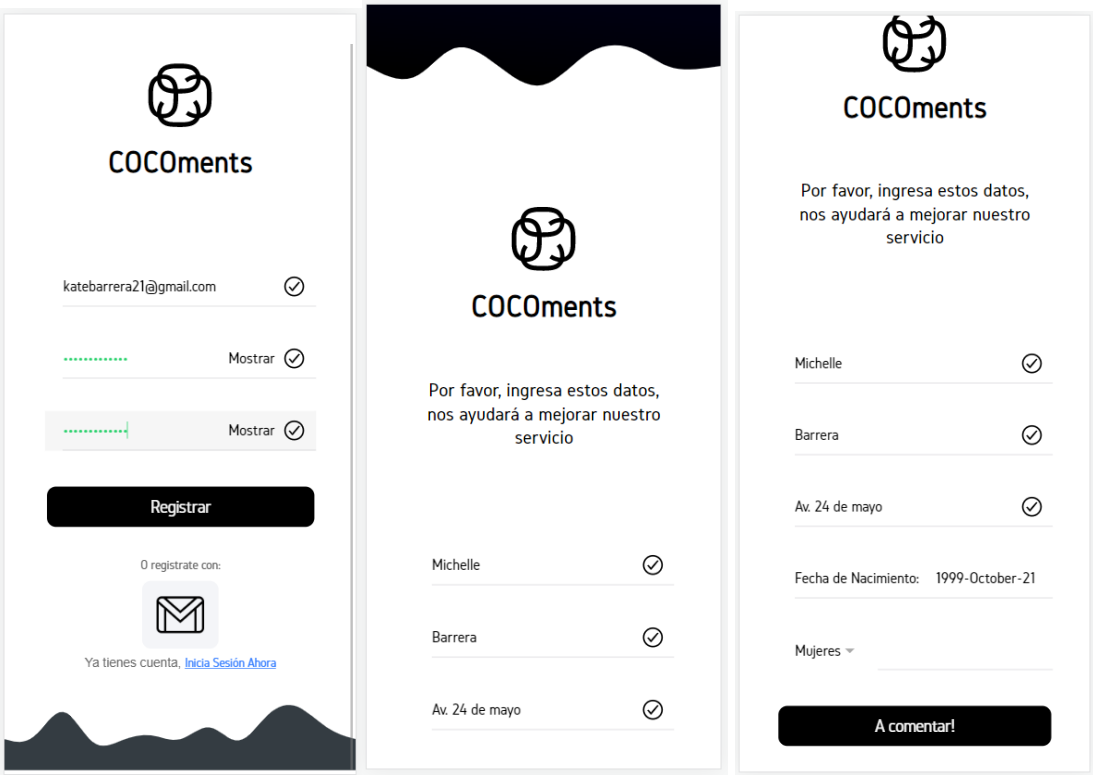


Tabla 19. Prueba unitaria para creación de comentario de usuario

Prueba funcional 3	Código	PF-0	Resultado
	Responsable	KB, RH	Aprobada
	Fecha	13/07/2022	

Requerimiento:

RF-03 Creación de comentario de usuario cliente

Resultado Esperado:

El usuario ingresará a la aplicación después de iniciar sesión, y navega a la página para agregar comentario. Ingresa un comentario en la caja de texto y presiona el botón guardar, se muestra una notificación mientras realiza este proceso y finalmente recibe una notificación que muestra que el comentario ha sido guardado con éxito.

Resultado Obtenido:

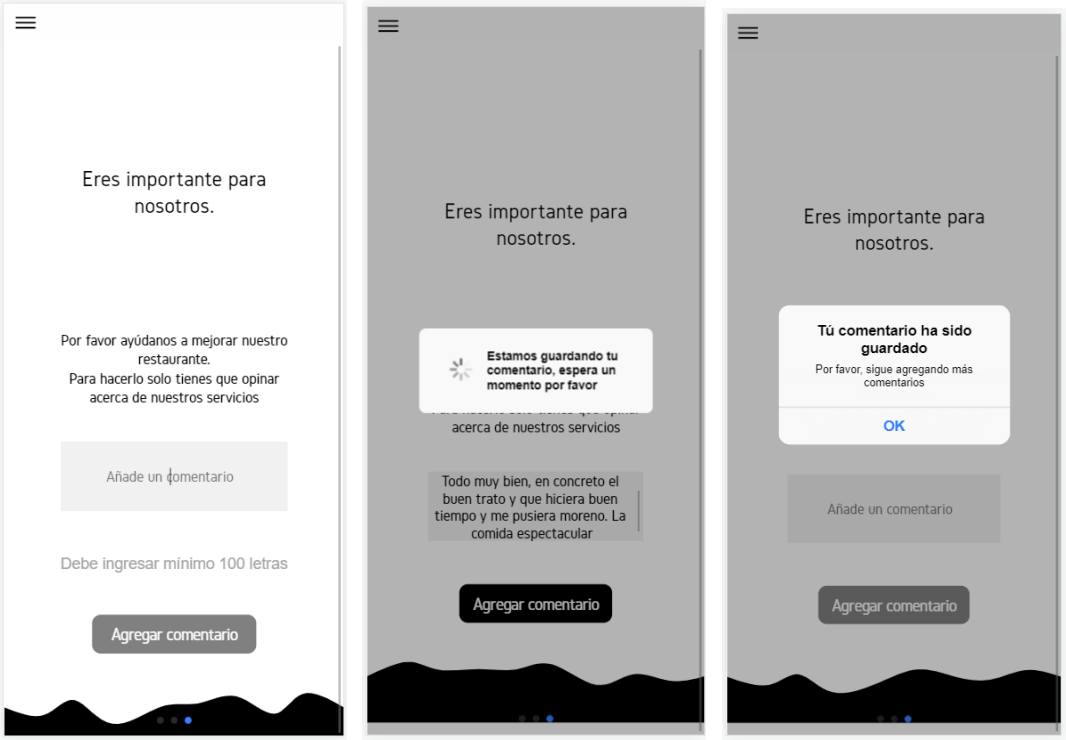


Tabla 20. Prueba unitaria para listar comentarios de usuario cliente

Prueba funcional 4	Código	PF-0	Resultado
	Responsable	KB, RH	Aprobado
	Fecha	13/07/2022	

Requerimiento:

RF-03 Listar comentarios del usuario/cliente

Resultado Esperado:

La página listar comentario debe presentar todos los comentarios que ha realizado el usuario dentro de a la aplicación, como el resumen de comentario, el tipo de sentimiento y la fecha.

Resultado Obtenido:

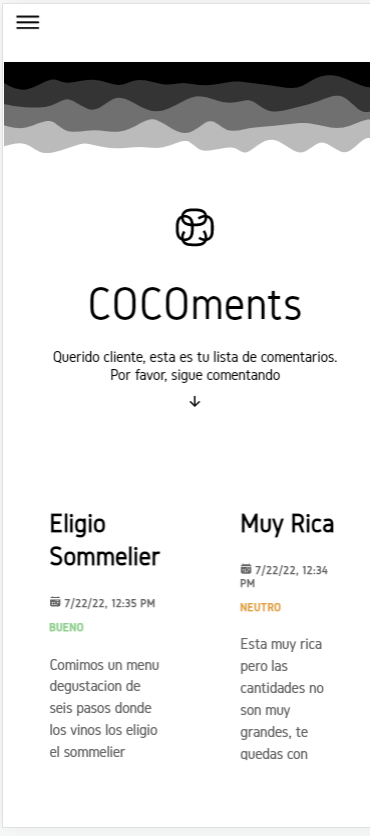


Tabla 21. Prueba unitaria para listar comentarios del administrador

Prueba funcional 5	Código	PF-0	Resultado
	Responsable	KB, RH	Aprobada
	Fecha	13/07/2022	

Requerimiento:

RF-06 Lista de comentarios para administrador

Resultado Esperado:

El administrador inicia sesión y navega a la página listar comentarios. Al ingresar obtiene una lista de todos los comentarios existentes dentro de la base de datos, además puede revisar la información completa al presionar sobre cada comentario. Puede hacer uso de los filtros como fecha y tipo

Resultado Obtenido:

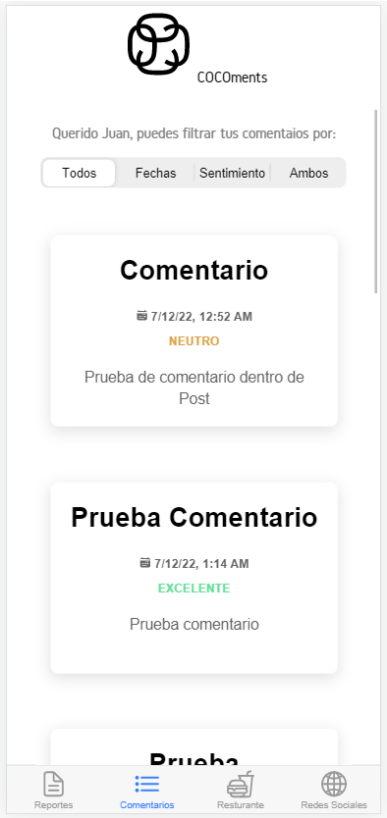


Tabla 22. Prueba unitaria para generación de reportes de comentarios para el administrador

Prueba funcional 6	Código	PF-0	Resultado
	Responsable	KB, RH	Aprobado
	Fecha	13/07/2022	

Requerimiento:

RF-07 Reporte de comentarios para administrador

Resultado Esperado:

El administrador inicia sesión y navega hacia la página de reportes. Se presentan varias gráficas como un histograma, nube de palabras, entre otras.

Resultado Obtenido:

Tabla 23. Prueba unitaria de recolección de datos de redes sociales

Prueba funcional 7	Código	PF-0	Resultado
	Responsable	KB, RH	Aprobado
	Fecha	13/07/2022	

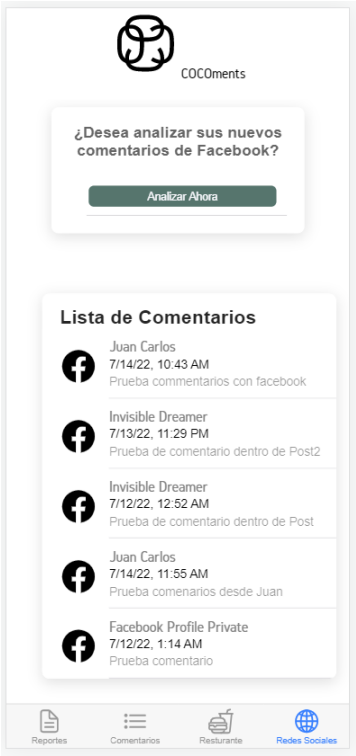
Requerimiento:

RF-08 Recolección de datos de redes sociales

Resultado Esperado:

Usar la API de la red social Facebook, realizar una publicación y obtener los comentarios existentes de ese comentario y guardarlo en la base de datos.

Resultado Obtenido:



7. Cronograma

Objetivo de aprendizaje	Actividad	Responsable	Duración	Comienzo	Fin
	Proyecto		616	12/04/2022	26/07/2022
	Sprint 1		264	12/04/2022	23/05/2022
OE.1.			8	12/04/2022	13/04/2022
	ACT.1. Especificación de requerimientos funcionales y no funcionales.	JB, KB	8	12/04/2022	13/04/2022
OE.2.			136	16/04/2022	07/05/2022
	ACT.1. Estudio de los fundamentos de Deep Learning	JB, KB	24	16/04/2022	19/04/2022
	ACT.2. Estudio de los fundamentos de procesamiento de lenguaje natural (PLN)	JB, KB	56	19/04/2022	27/04/2022
	ACT.3. Estudio de herramientas para extracción de datos	JB, KB	56	28/04/2022	07/05/2022
OE.3.			116	8/05/2022	
	ACT.1. Proceso de recolección de datos	JB, KB	12	8/05/2022	10/05/2022
	ACT.2. Preparación y limpieza de datos	JB, KB	24	11/05/2022	13/05/2022
	ACT.3. Diseño y desarrollo del modelo aplicado	JB, KB	40	14/05/2022	18/05/2022
	ACT.4. Diseño y desarrollo para representación de datos	JB, KB	40	19/05/2022	23/05/2022
	ACT.ER. Escritura y revisión de evidencias		4	23/05/2022	23/05/2022
	Sprint 2		204	24/05/2022	04/07/2022
OE.1.			32	24/05/2022	26/05/2022
	ACT.4. Estudio de los fundamentos de Framework para desarrollo móvil y web como IONIC y Angular.	JB, KB	32	24/05/2022	26/04/2022

OE.3			168	27/04/2022	01/04/2022
	ACT.2. Diseño y desarrollo del módulo Usuario.	JB	16	27/04/2022	30/04/2022
	ACT.3. Diseño y desarrollo del módulo Administrador	KB	16	30/04/2022	31/04/2022
	ACT.4. Diseño y desarrollo del módulo de Reportes	JB, KB	48	01/05/2022	10/04/2022
	ACT.5. Diseño y desarrollo del módulo de Comentarios	KB	40	13/05/2022	20/05/2022
	ACT.6. Diseño y desarrollo del módulo de Establecimiento	JB	8	23/05/2022	24/05/2022
	ACT.7. Desarrollo del módulo Autentificación de usuario a través de Firebase	JB, KB	24	25/05/2022	27/05/2022
	ACT.9. Diseño y desarrollo de la base de datos dentro de Firebase	JB, KB	16	02/07/2022	03/07/2022
ACT.ER. Escritura y revisión de evidencias		JB, KB	4	04/07/2022	04/07/2022
	Sprint 3		16	05/07/2022	07/07/2022
OE.4.			16	05/07/2022	07/07/2022
	ACT.1. Integración del servicio de API REST para agregar nuevo comentario realizado por el usuario	KB	8	05/07/2022	07/07/2022
	ACT.2. Integración del servicio de API REST para el despliegue del reporte para el usuario administrador	JB	8	05/07/2022	07/07/2022
	Sprint 4		128	08/07/2022	23/07/2022
OE.5			128	08/07/2022	23/07/2022
	CT.1. Diseño y desarrollo del plan de pruebas funcionales y no funcionales	JB, KB	4	08/07/2022	08/07/2022
	ACT.8. Integración del módulo de adquisición de datos de redes sociales	JB, KB	64	09/07/2022	20/07/2022
	ACT.8. Diseño y desarrollo del informe escrito	JB, KB	52	11/07/2022	22/07/2022
ACT.ER. Escritura y revisión de evidencias		JB, KB	8	23/07/2022	23/07/2022

Total, de horas	612
Horas por Juan Carlos Barrera Barrera	308
Horas por Katherine Michelle Barrera	308
Fecha de inicio	12/04/2022
Fecha de finalización	26/07/2022

8. Presupuesto

DENOMINACIÓN	CANT.	COSTO UNITARIO	COSTO TOTAL
	Unidades	Dólares	Dólares
1. Tecnológico			
Computadora portátil	2	1000.00	2000.00
Celular	2	550.00	1100.00
		Subtotal	$ 3100.00
2. Bienes			
Copias	128	0.03	3.84
Impresiones	200	0.05	10.00
Empastados	2	15.00	20.00
		Subtotal	$ 33.84
3. Servicios costo por mes de uso			
Servicios de Internet	2	23.00	46.00
Servicio de Cloud Firestore – Almacenamiento – GB	138	-	24.66
Servicio de Cloud Firestore – Escritura – operaciones	8800000	-	14.76
Servicio de Cloud Firestore – Lectura – operaciones	5300000	-	2.28
Servicio de Cloud Firestore – Eliminación – operaciones	4400000	-	0.76
Servicio de Cloud Storage – GB Almacenados – fotos de HD	5000	-	0.13
Servicio de Google Colan Pro	1	9.99	9.99
Google Drive almacenamiento business starter	1	5.40	5.40
		Subtotal	$103.98
4. API de pago			
Pago por licencia de API Multilingual sentiment analysis	1	19.00	19.00
		Subtotal	$ 19.00
5. Alimentación			
Alimentación de estudiantes – días	75	4.00	300.00
		Subtotal	$ 300.00
4. Personal			
Estudiante/Desarrollador 1 – horas	300	2.75	825.00
Estudiante/Desarrollador 2 – horas	300	2.75	825.00
Asesoría Especializada – ingeniero en computación	1	1200	1200.00
		Subtotal	$ 2850.00
5. Otros			
Imprevistos	1	100.00	100.00
		Subtotal	$ 100.00
Total			**$ 6506.82**

9. Conclusiones

En esta sección se explican las conclusiones que se obtuvieron durante el proceso de desarrollo de la aplicación para resumen de comentarios. Se ordena desde una conclusión general, de aplicación, modelo y finalmente una conclusión por cada objetivo propuesto.

Conclusión general

La aplicación fue realizada mediante un proceso de metodología ágil que ayudo a gestionar correctamente la realización de este. De modo que, se entregaron los Sprints establecidos con anterioridad, a mismo tiempo la toma de requerimientos es esenciales para este tipo de trabajo. La aplicación presenta varios módulos para satisfacer la necesidad del cliente, de igual manera la gestión de usuarios. Se lograron obtener los resúmenes de cada comentario ingresado por el usuario, al igual que la obtención de comentarios desde la red social Facebook.

Conclusiones de Aplicación

En cuanto al desarrollo de la aplicación hibrida, fue lograda con éxito. Después de tomar los requerimientos funcionales, se realizaron los prototipos para la aplicación que nos ayudó a obtener una idea más concisa tanto para el cliente como para desarrollados establecer que se buscaba obtener con la aplicación.

Una vez aprobados los prototipos se realizó la parte de Front-End aplicando un diseño minimalista debido a que esto facilitaría al usuario su navegación y acorta el tiempo de aprendizaje de la aplicación. Se tuvo en cuenta a todos los elementos de la interfaz debido a que el diseño es responsivo y debe ser fácil de visualizar en varias plataformas.

En cuanto al Back-End se realizaron de manera conjunta, debido a la metodología aplicada. Para verificación de consultas, primero se usó Postman para realizar la consulta mediante la URL y los parámetros necesarios, obteniendo resultados altamente eficientes.

Conclusiones del Modelo

A través de nuestro modelo hemos logrado utilizar un modelo pre—entrenado como lo es Trsanformer T5, a este modelo se realizó finetuning con el propósito de generar resúmenes de texto, y el resultado es que se tiene resúmenes claros y con un contexto excelente con respecto al texto de entrada, esto se puede observar dentro de las medidas de calidad del algoritmo, teniendo una media de 0.36 entre Rouge 1, 2 y L, esto es una

buena medida, ya que dentro del modelo se intenta obtener resumes abstractivos, y no una copia de una oración dentro del texto y con el modelo generado se logra ese propósito. El modelo es evaluado con un conjunto de datos propio con 200 registros en donde se tiene el texto de entrada y su resumen, estos resúmenes fueron generados por humanos y los resultados obtenidos con el modelo en la mayoría de los casos son maravillosos. El modelo llega a fallar cuando el texto de entrada es demasiado corto, por lo que simplemente copiara la oración, para un buen resumen la cantidad mínima de caracteres de entrada es de 100, el modelo genera un resumen de 20 caracteres máximo, debido a que los resúmenes dentro del dataset no superaban esta cantidad, todo esto está enfocado a nuestro objetivo. Sin embargo, si se desea que el resumen sea de mayor tamaño se deberá entrenar nuevamente el modelo con resúmenes mucho más grandes y obtendrá los resultados que está buscando.

Conclusión de Objetivos

Se realizó un correcto estudio sobre los fundamentos de Deep Learning que se debió seguir para poder realizar el presente proyecto. De igual manera con el procesamiento de lenguaje natural, y tecnologías como IONIC, Angular, Django y las herramientas existentes de extracción de datos con el fin de poder comprender el proceso de desarrollo para en lo posterior poder realizarlo.

Se diseñó y desarrollo un modelo de análisis de datos en fases para el resumen de texto, teniendo tres fases, la primera destina a la limpieza del dataset y análisis exploratorio, la segunda está enfocada al entrenamiento y generación del modelo; mientras que en la última fase se obtiene el método de predicción para la generación de resúmenes con cualquier entrada de texto recibida. Para la representación de comentarios se hace uso de distintas técnicas de procesamiento de lenguaje natural por medio de nube de palabras, ngramas, el resumen como tal también es una representación, entre otros.

En cuanto al diseño y desarrollo de la aplicación hibrida, se realizó con éxito, aunque se debe recalcar que al usar la última versión de IONIC, se encontraron muchos problemas debido a la mala documentación existente para la versión 6. Por lo que esto presento varios retrasos y un tiempo más extenso que el estimado para realizar actividades como autentificación y el desarrollo de otras funciones.

En cuanto al módulo de adquisición de redes sociales, se usó la aplicación de Facebook de forma bastante exitosa. Se logró obtener las últimas 25 publicaciones de las cuales se almacenan los comentarios de cada publicación obtenida, para implementar una categorización, se realizó mediante el encabezado de la publicación que tendría especificado la categoría y producto a la que pertenece y así para poder guardar los datos de forma ordenada en nuestra base de datos.

Se logró integrar el modelo de análisis para resumen de texto dentro de la aplicación hibrida, para esto se generaron varios archivos y directorios en donde está ubicado el modelo, estos recursos contienen, los pesos, configuración y tokenizador del modelo que ya ha sido entrenado. Estos archivos son utilizando dentro de Django, para obtener la

configuración de la red transfomer y poder realizar los resúmenes de nuestros comentarios.

Finalmente se realizaron pruebas unitarias en base a los requerimientos funcionales en el cual sé estable lo que se espera de la aplicación y el resultado obtenido después de su desarrollo siendo en su mayoría altamente aceptable, para nuestro Product Owner y para el equipo de desarrollo.

Trabajo Futuro

En esta sección se tratará sobre el trabajo futuro, que puede servir en lo posterior a esta aplicación. Debido al tiempo que se tuvo para desarrollar la aplicación se ha optado por realizar la primera fase de este proyecto, por lo que los trabajos futuros son extensos en este campo.

Para la parte del usuario, se cree necesario como trabajo futuro el desarrollo de una interfaz con mayor interacción puede ser mediante juegos, adivinanzas, entre otros, estos enfocados al negocio, dentro de la aplicación se puede implementar una página de inicio en donde se llegue a observar los comentarios sobre los distintivos servicios del establecimiento; a fin de que el usuario mantenga el interés en realizar comentarios dentro de la aplicación. La aplicación puede escalar a varios establecimientos, y así obtener mayor cantidad de conocimiento acerca de los distintos negocios involucrados, esto permite a los usuarios comentar sobre distintos negocios sobre qué le parece el servicio. Toda esta información es valiosa porque al ser analizada puede dar mejores estrategias de ventas para los establecimientos. Dentro del proyecto se trabajó con restaurantes/bares en donde se planteó cuáles son los mejores y peores platos del establecimiento, sin embargo, esto puede escalar a muchas más opciones, ya sea analizar que opinan acerca del personal, la ubicación, que productos son los más recomendados del negocio, entre otros.

Debido al gran alcance que tiene la aplicación, se puede considerar como trabajo futuro, el seguir mejorando el análisis de datos como el análisis de sentimientos para que la precisión sea mejor, por ahora se trabajó con una API pagada, sin embargo, al tener nuestra información clasificada podemos realizar este análisis de sentimientos y ya no depender de esta, también se puede realizar clusters para una categorización de los productos de manera más rápida. Al igual que ofrecer un resumen más extenso como resultado, esto se puede realizar aplicando un mayor conjunto de datos de entrenamiento con mejores resúmenes, estos resúmenes ya no solo enfocados a los comentarios sino también el resumen general de todos los resúmenes. También se plantea generar reportes estáticos para impresión de información relevantes.

De igual manera, como trabajo futuro, se puede estudiar la transformación de contenido como emojis, gifs, stickers usados generalmente por el usuario para obtener un análisis de sentimientos.

Dentro del modelo también hay trabajo futuro pensando en generar un modelo que llegue a comportarse como un diálogo, con esto nos referimos a implementar un bot, que no solo

sepa dar secuencia de instrucciones estructuradas, sino que pueda llegar a mantener un diaálogo. Esto con el fin de que el usuario al momento de ingresar un comentario tenga la perspectiva de estar conversando con alguien, y este pueda llegar a ofrecer que productos son los que le favorecen. Actualmente existe GTP3 que nos ayuda con esta tarea. Dejamos abierto este tema para que pueda realizar una búsqueda sobre cómo funciona esta tecnología, ya que cuenta con muchas herramientas para trabajo con texto.

10. Recomendaciones

En esta sección se explican las recomendaciones que se pueden brindar después del desarrollo de la aplicación hibrida. Se ordenan desde recomendaciones a nivel de desarrollo, a nivel de análisis de datos que mejora el rendimiento del modelo y finalmente recomendaciones para generar nuevas gráficas.

Recomendaciones a nivel de Desarrollo

Se recomienda mejorar la seguridad para la aplicación, tanto en el control de usuarios, como en los permisos sobre los mismos. Se recomienda trabajar de forma más cercana con el usuario, dueño del restaurante a fin de supervisar, mejorar y encontrar nuevas gráficas que sean útiles para el desarrollo de su negocio. Se recomienda también definir si se trabajara como aplicación móvil o web, ya que esto toma mucho tiempo, hacer que una aplicación para varios dispositivos y que el resultado del diseño sea lo mejor posible.

Recomendaciones a nivel de Análisis de Datos

Dentro del desarrollo del modelo, se recomienda tener un dataset más grande que con el que se trabajó, el modelo fue entrenado con 200 registros con el texto y resumen correspondiente, si se desea mejorar la precisión y la extensión del resumen debería contar con esta información. Una vez que la tenga es cuestión de reentrenar el modelo con su información y obtendrá los resultados que está esperando.

Recomendaciones para mejorar el rendimiento del modelo

Para el entrenamiento del modelo, se recomienda trabajar con GPU, o se puede trabajar utilizando Google COLAB, que le brinda esta facilidad para que entrene su modelo.

Como anteriormente se menciona, el modelo, trabajara de manera mucho para rápida con GPU, por lo que si dispone de esta puede habilitar al momento de la predicción del resumen, GPU y eso acelera el proceso de resumen de comentario.

Recomendaciones para generar nuevas gráficas

Para una mayor facilidad de generación de gráficas puede leer la documentación de Apexcharts, ya que disponen de una gran cantidad de ejemplos.

Referencias Bibliográficas

Acosta, M., Yagual Velastegui, A., & Coronel Pérez, V. (2018). PERSPECTIVAS DE LA ECONOMÍA DIGITAL EN LATINOAMÉRICA: CASO ECUADOR. *EBSCOhost*, 29-43.

Alammar, J. (2018). *The Illustrated Transformer [Blog post]*. Obtenido de https://jalammar.github.io/illustrated-transformer/

Arora, P., Virmani, D., Jindal, H., & Sharma, M. (2017). Sorted K-Means Towards the Enhancement of K-Means to Form Stable Clusters. En *Proceedings of International Conference on Communication and Networks* (págs. 479–486). Springer, Singapore.

Ashley, D. (2020). Foundation Dynamic Web Pages with Python. En *Django* (págs. 183–199). Berkeley, CA: Apress.

Ayyadevara, V. K. (2018). Random Forest. En *Pro Machine Learning Algorithms* (págs. 105-116). APRESS.

Bisong, E. (2019). Principal Component Analysis (PCA). En *Building Machine Learning and Deep Learning Models on Google Cloud Platform* (págs. 319–324). Apress, Berkeley, CA.

Buenaño, D. E., Barriga, V. A., & Cadena, V. E. (2015). Usuarios, clientes y consumidores digitales: consideraciones para su abordaje desde el marketing y la comunicación corporativa. *Redmarka: revista académica de marketing aplicado*, 3-19.

CANTOS, & MOLINER, C. (2001). Calidad de Servicio y Satisfacción del cliente. *Revista de Psicología del Trabajo y de las Organizaciones*, 233-235.

Chaudhary, P. (2018). IONIC FRAMEWORK . *International Research Journal of Engineering and Technology (IRJET)*, 5.

Cho, K., Merrienboer, B. v., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation.* arXiv preprint arXiv.

Coba, G. (2020). Las cocinas fantasma crecen en medio de la 'coronacrisis'. *Primicias*.

Cortez Vásquez, A., Vega Huerta, H., & Pariona Quispe, J. (2009). Procesamiento de lenguaje natural. Obtenido de https://revistas.gnbit.net/index.php/sistem/article/view/5923/5121

Del-Alcázar, J. P. (2021). Ecuador Estado Digital Ene/21. *Menthino, Quito*.

DiPietro, R. B., Crews, T. B., Gustafson, C., & Strick, S. (2012). Journal of Foodservice Business Research. En *The Use of Social Networking Sites in the Restaurant Industry: Best Practices* (págs. 265-284).

Espinoza Vélez, M. P., & Armijos Buitrón, V. A. (2022). La transformación digital y su incidencia en el e-commerce en Ecuador. *Memorias de la Décima Segunda Conferencia Iberoamericana de Complejidad, Informática y Cibernética, CICIC*.

Etemad, A. G., Abidi, A. I., & Chhabra, M. (2021). *Fine-Tuned T5 for Abstractive Summarization.* Greater Noida, India: Totem Publisher, Inc. All rights reserved.

Fowler, F. M. (2018). Navigating Hybrid Scrum Environments. En *The Scrum Master* (págs. 47–52). Sunnyvale, CA, USA: Apress.

Fowler, F. M. (2018). The Scrum Team. En *Navigating Hybrid Scrum Environments* (págs. 25–30). Sunnyvale, CA, USA: Apress.

Gagliardi, V. (2021). Decoupled Django. En *Modern Django and the Django REST Framework* (págs. 31–40). Berkeley, CA: Apress.

Garg, A., Adusumilli, S., Yenneti, S., Badal, T., d. g., Pandey, V., . . . Agarwal, R. (2021). NEWS Article Summarization with Pretrained Transformer. En I. 2020, *International Advanced Computing Conference* (págs. 203–211). India: 2021 Springer Nature Singapore Pte Ltd.

Gartner. (2016). Cuadrante Mágico para Plataformas de Business Intelligence y Analytics .

González, P., & Manuel, J. (2014). Desarrollo de una API para la descripcion y gestion de servicios WEb REST. Obtenido de http://repositori.uji.es/xmlui/bitstream/handle/10234/156006/TFM_2014_puertaJ.pdf?sequence=1&isAllowed=y

Hajian, M. (2019). Deploying to Firebase as the Back End. En *Progressive Web Apps with Angular* (págs. 9–27). Apress, Berkeley, CA.

Jeffrey, J., Luca, T., M. P., Emanuela, M., & Riccardo. (2018). Big Data: negocios, tecnología, educación y ciencia. *ACM, Association for Computing Machinery*.

Kelly, A. (2019). The Art of Agile Product Ownership. En *Scrum and the Product Owner* (págs. 53–55). Berkeley, CA: Apress.

Kim, K., Aminanto, M. E., & Tanuwidjaja, H. C. (2018). Network Intrusion Detection using Deep Learning. En S. o. Networks, *Deep Learning* (págs. 27–34). Singapore: Springer.

Kramer, O. (2013). K-Nearest Neighbors. En *Dimensionality Reduction with Unsupervised Nearest Neighbors* (págs. 13-23). Springer, Berlin, Heidelberg.

Kristoffersen, E., Aremu, O. O., Blomsma, F., Mikalef, P., & Li, J. (2019). Exploring the Relationship Between Data Science and Circular Economy: An Enhanced CRISP-DM Process Model. En *Conference on e-Business, e-Services and e-Society: Digital Transformation for a Sustainable Society in the 21st Century* (págs. 177–189). Springer, Cham.

Kurniasari, & Setyanto. (2020). *Sentiment Analysis using Recurrent Neural Network.* Yogyakarta, Indonesia: IOP Publishing.

Lavanya, C. (2021). Artificial Intelligence. En D. J. KARTHIKEYAN, D. T. HIE, & D. N. JIN, *LEARNING OUTCOMES OF CLASSROOM RESEARCH* (págs. 12-19). India: L ORDINE NUOVO PUBLICATION.

Le, J. (18 de 08 de 2016). *KD nuggets*. Obtenido de The 10 Algorithms Machine Learning Engineers Need to Know: https://www.kdnuggets.com/2016/08/10-algorithms-machine-learning-engineers.html

Manavi, M., & Zhang, Y. (2019). A New Intrusion Detection System Based on Gated Recurrent Unit (GRU) and Genetic Algorithm. En *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage: Security, Privacy, and Anonymity in Computation, Communication, and Storage* (págs. 368–383). Springer, Cham.

Mariño, S. I., & Alfonzo, P. L. (2014). *Implementación de SCRUM en el diseño del proyecto del Trabajo Final de Aplicación.* Pereira, Colombia: Universidad Tecnológica de Pereira.

McCarthy, J. (2000). Concepts of Logical AI. En *Logic-Based Artificial Intelligence* (págs. 37–56). Springer, Boston, MA.

McCarthy, J. (12 de 11 de 2007). *WHAT IS ARTIFICIAL INTELLIGENCE?* Stanford University: Stanford, CA. Obtenido de WHAT IS ARTIFICIAL INTELLIGENCE?: http://faculty.otterbein.edu/dstucki/inst4200/whatisai.pdf

Menaka, R., Thaker, J. K., Bhushan, R., & Karthik, R. (2020). IMEXT Text Summarizer Using

Deep Learning. En *Applied Computer Vision and Image Processing* (págs. 34–45). Springer, Singapore.

MercadoLibre. (2020). «El comportamiento del consumidor en tiempos de crisis: Previo a la cuarenta COVID 19.

Ministerio de Turismo de Ecuador. (2019). Turismo en Ecuador Alojamiento y servicios de comida.

Ministerio de Turismo, Ecuador. (11 de 05 de 2016). MINTUR constituye "Asociación de Bares, Restaurantes y de Entretenimiento de Cuenca". *Gobierno de la República del Ecuador*.

Moore, J. (4 de 11 de 2019). *An Introduction to Sliding Window Algorithms*. Obtenido de Level Up Coding: https://levelup.gitconnected.com/an-introduction-to-sliding-window-algorithms-5533c4fe1cc7

Mridha, M. F., Hamid, M. A., & Asaduzzaman, M. (2019). Issues of Internet of Things (IoT) and an Intrusion Detection System for IoT Using Machine Learning Paradigm. En *Proceedings of International Joint Conference on Computational Intelligence* (págs. 395–406). Springer, Singapore.

Nabati, E. G., & Thoben, K.-D. (2017). On Applicability of Big Data Analytics in the Closed-Loop Product Lifecycle: Integration of CRISP-DM Standard. En *IFIP International Conference on Product Lifecycle Management: Product Lifecycle Management for Digital Transformation of Industries* (págs. 457–467). Springer, Cham.

Naing, H. W., Thwe, P., Mon, A. C., & Naw, N. (2018). Analyzing Sentiment Level of Social Media Data Based on SVM and Naïve Bayes Algorithms. En *Big Data Analysis and Deep Learning Applications. ICBDL 2018. Advances in Intelligent Systems and Computing* (págs. 68-76). Springer Nature Singapore Pte Ltd.

Oracle. (2015). *Wall Street Journal. Dominio de los datos: el impulsor global de los ingresos*. Obtenido de https://www.oracle.com/us/dm/global-executive-study-2618506.pdf

Ortiz, I. R. (2022). *DESARROLLO Y OPTIMIZACIÓN DE UN MÉTODO DE CLASIFICACIÓN MEDIANTE REDES NEURONALES CON KERAS PARA APROBACIÓN DE UN CRÉDITO EN UNA INSTITUCIÓN FINANCIERA. USO DE PIPELINE, CROSS_VAL_SCORE, METRICS Y GRIDSEARCHCV. VISUALIZACIÓN DE RESULTADOS EN UN HEATMAP*. Obtenido de efecode.com: https://efecode.com/desarrollo-optimizacion-redes-neuronales-keras-aprobacion-credito-institucion-financiera-pipeline-gridsearchcv-visualizacion-heatmap

Paez, G., & Monroy. (2020). *Implementación de un modelo análisis de sentimientos con respecto a la JEP basado en minería de datos en Twitter.* Bogotá D.C.: Universidad Católica de Colombia.

Penkler, D. (2018). Grandes datos. Retos tecnológicos y empresariales del Big Data en la economía digital. *Association for Computing Machinery*.

PINEROS, & BLANCO, J. (2009). *Medición de la satisfacción del cliente del restaurante museo taurino, y formulación de estrategias de servicio para la creación de valor.* Bogotá: Pontificia Universidad Javeriana.

Pte, S. N. (2021). Advances in Applications of Data-Driven Computing. En S. B. Ankush Ghosh, *Comparison and Analysis of RNN-LSTMs and CNNs for Social Reviews Classification* (págs. 49–59). Singapore: Springer.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., . . . Liu, P. J. (2020). *Exploring the Limits of Transfer Learning with a Unified.* Google, Mountain View, CA

94043, USA.
Reinoso Naranjo, B. A. (2021). *Análisis de la restauración en Quito - Ecuador tras el impacto de la COVID -19.* Quito: Universidad de los Hemisferios.
Ruchi, S., Yashaswi, A., & Dr. Jongwook, W. (2017). Análisis de Big Data de empresas locales y reseñas. *Conferencia Internacional sobre Comercio Electrónico 2017*.
Salas, R. (2018). *Redes Neuronales Artificiales.* Obtenido de https://d1wqtxts1xzle7.cloudfront.net/50358783/Redes_Neuronales_Artificiales-with-cover-page-v2.pdf?Expires=1633017707&Signature=HtWUstUr2OboZ6dtt9XjrIG67TEdqqPj7v4HatnGW-jNMYOqOpHE5VDuc4fNlXTVLv70rjTgFzUeORAXDTUkC0vrQsAS5jqot0TUWZIIIrgC1oIn0uqjpqNUo4WRkv
Seif, G. (05 de 02 de 2018). *The 5 Clustering Algorithms Data Scientists Need to Know*. Obtenido de Towards Data Science: https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68
Sewak, M. (2019). Temporal Difference Learning, SARSA, and Q-Learning. En *Deep Reinforcement Learning* (págs. 51–63). Springer, Singapore.
Sharda, R., Delen, D., & Turba, E. (s.f.). Business Intelligence and Analytics: Systems for Decision Support . *MA:Pearson*, 2018.
Sharma, G., Kumar, P., & Shrivastava, L. (2020). An Efficient Performance of Enhanced Bellman-Ford Algorithm in Wireless Sensor Network Using K-Medoid Clustering. En *International Conference on Sustainable and Innovative Solutions for Current Challenges in Engineering & Technology ICSISCET 2019: Intelligent Computing Applications for Sustainable Real-World Systems* (págs. 52–65). Springer, Cham.
Silva, A. M., & Leong, P. H. (2015). Feature Selection. En *Grammar-Based Feature Generation for Time-Series Prediction* (págs. 13–24). Springer, Singapore.
Silvana Dakduk, P., & Ricardo Dicarlo, M. (2020). Situación de las empresas durante el COVID-19 Ecuador. *Cámara Ecuatoriana de Comercio Electrónico, Universidad Espíritu Santo y Banco del Pacífico*.
Singh, P. (2018). Logistic Regression. En *Machine Learning with PySpark* (págs. 65–98). Apress, Berkeley, CA.
Singha, C., Bilal, J., & Wornchanok, C. (2019). Aparición de un nuevo entorno empresarial con Big Data e Inteligencia Artificial.
Sol, Z., & Narasimha, R. V. (2021). Análisis responsable de Big Data para servicios de comercio electrónico.
StatistaResearchDepartment. (2021). Ventas del comercio electrónico minorista en América Latina de 2019 a 2023.
Trendafilov, N., & Gallo, M. (2021). Linear discriminant analysis. En *Multivariate Data Analysis on Matrix Manifolds* (págs. 229–268). Springer, Cham.
Trigas Gallegos, M. (s.f.). Metodología Scrum. Obtenido de http://openaccess.uoc.edu/webapps/o2/bitstream/10609/17885/1/mtrigasTFC0612memoria.pdf
Ungerer, C., Portugal, A., Molinuevo, M., & Natasha, R. (2020). Recommendations to Leverage ECommerce during the Covid-19 Crisis. *WorldBank, DC*.
Uzayr, S. b., Cloud, N., & Ambler, T. (2019). JavaScript Frameworks for Modern Web Development. En *Angular* (págs. 209–223). Berkeley: Apress.

Vasudha Goswami, V. M. (2020). Detecting Spam Emails/SMS Using Naive Bayes, Support Vector Machine and Random Forest. En B. d. International conference on Computer Networks, *Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2019)* (págs. 306-313). Springer Nature Switzerland AG.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention is All you Need. En *Advances in Neural Information Processing Systems 30 (NIPS 2017)* (págs. 5998-6008).

Verma, N. K., & Salour, A. (2020). Feature Selection. En *Intelligent Condition Based Monitoring* (págs. 175–200). Springer, Singapore.

Wooditch, A., Johnson, N. J., Solymosi, R., Ariza, J. M., & Langton, S. (2021). Ordinary Least Squares Regression . En *A Beginner's Guide to Statistics for Criminology and Criminal Justice Using R* (págs. 245–268). Springer, Cham.

Z. Sun, L. S., & Strang, K. (2018). Servicios de análisis de macrodatos para mejorar la inteligencia empresarial. *Journal of Computer Information Systems, vol 58.*, 162-169.

Zaremba, W. (2015). *RECURRENT NEURAL NETWORK.* New York University. Obtenido de https://arxiv.org/pdf/1409.2329.pdf

Zhang, L., Wang, S., & Liu, B. (2018). *Deep learning for sentiment analysis: A survey.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.

Zhou, H. (2020). K-Nearest Neighbors. En *Learn Data Mining Through Excel* (págs. 93–108). Apress, Berkeley, CA.

Anexos

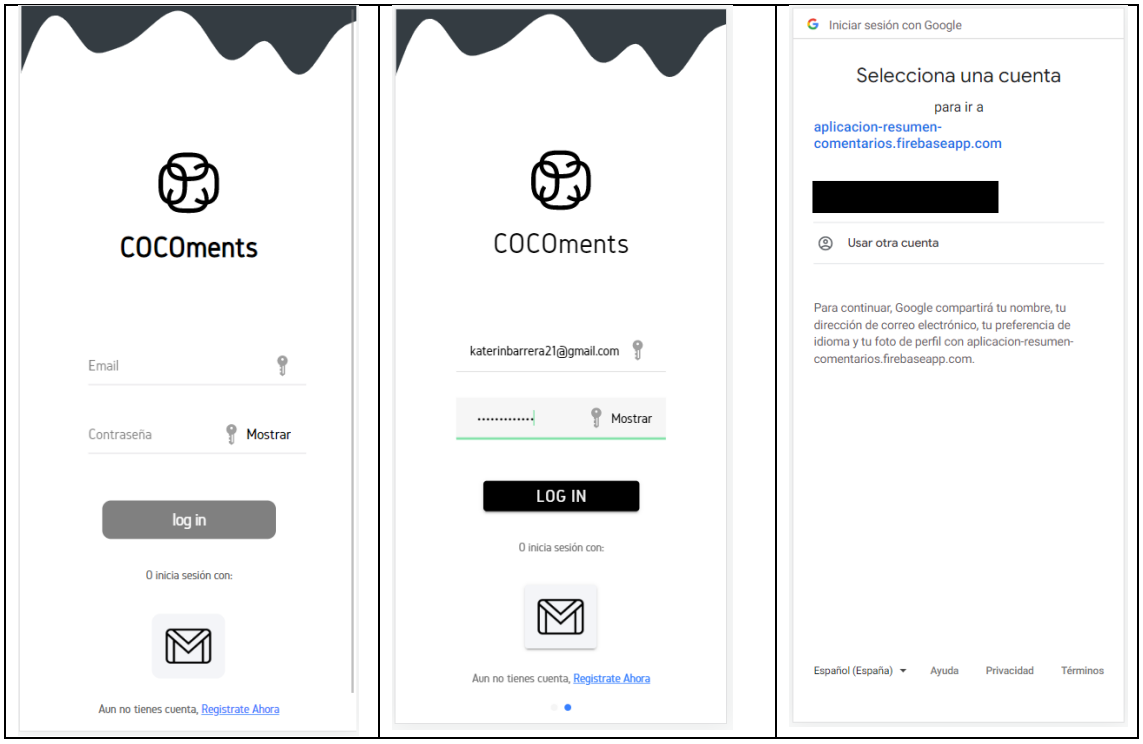
En esta sección se presentan anexos de la aplicación, separados en cinco secciones, la primera muestra la aplicación en un dispositivo móvil y el segundo en web. La tercera sección presenta gráficas de análisis estadístico generado en Django, la cuarta sección presenta un modelo RNN para resumen de texto y finalmente un apartado para la codificación con enlaces hacia los repositorios públicos.

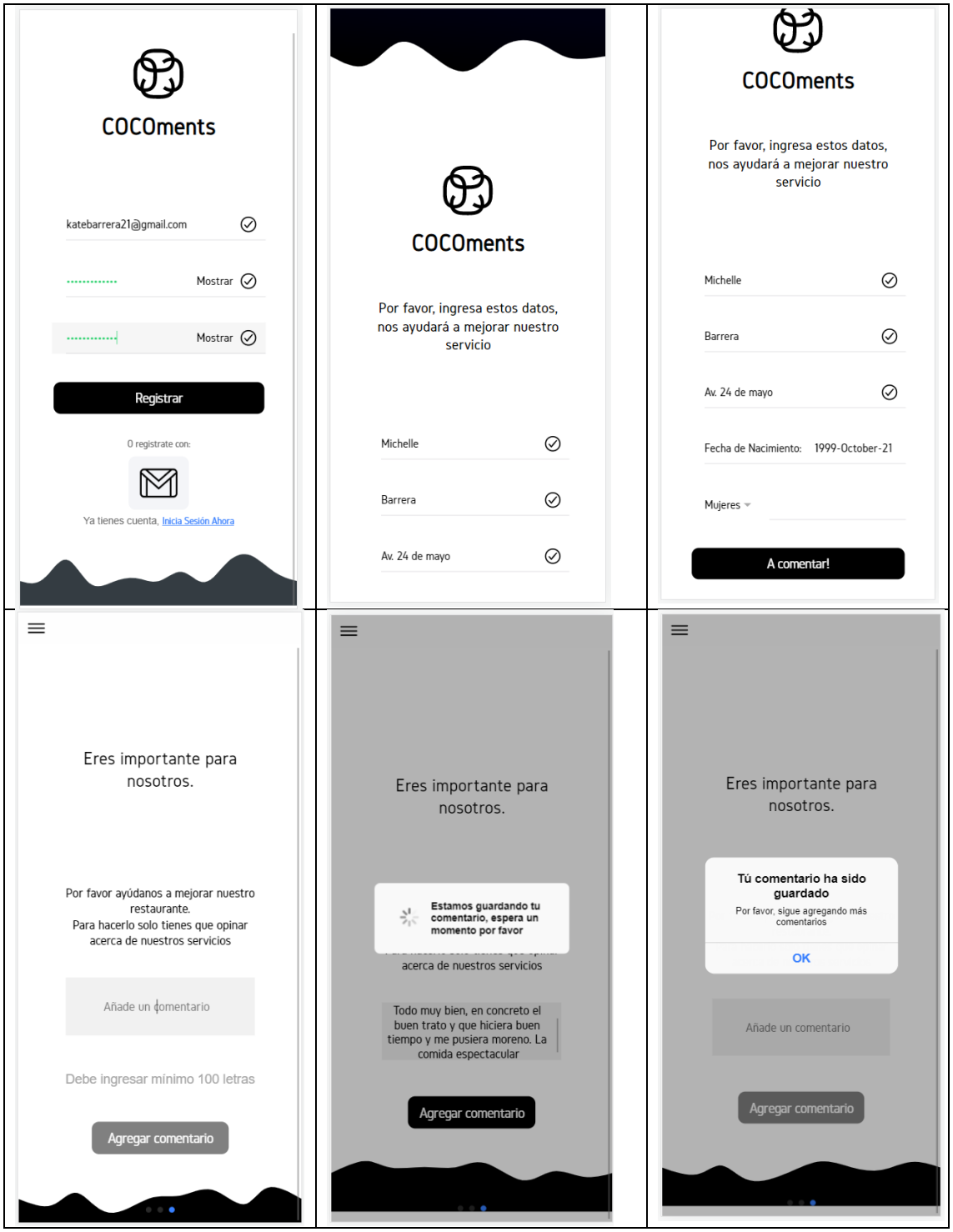
A. Aplicación Móvil

En este anexo se presenta capturas sobre el despliegue de la aplicación del módulo usuario y el módulo administrador sobre un dispositivo móvil.

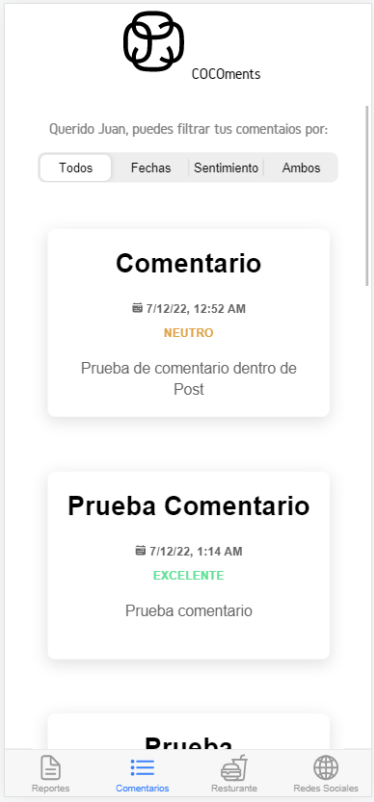
Usuario

En este apartado se presenta las capturas del despliegue de la aplicación del módulo usuario, primero tenemos la función de autentificación, segundo la función registrar, tercero la función realizar comentario y finalmente un listar comentario.



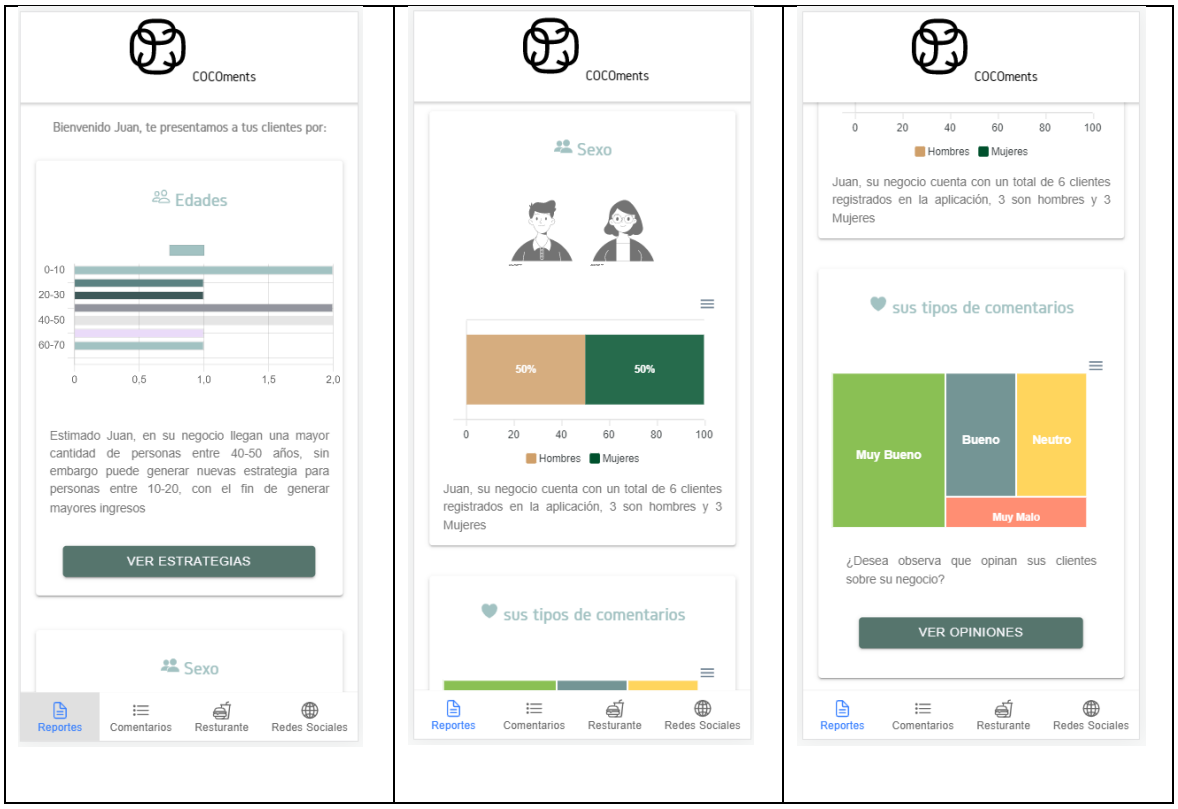

COCOments
katebarrera21@gmail.com
Mostrar
Mostrar
Registrar
O registrate con:
Ya tienes cuenta, Inicia Sesión Ahora
COCOments
Por favor, ingresa estos datos, nos ayudará a mejorar nuestro servicio
Michelle
Barrera
Av. 24 de mayo
COCOments
Por favor, ingresa estos datos, nos ayudará a mejorar nuestro servicio
Michelle
Barrera
Av. 24 de mayo
Fecha de Nacimiento: 1999-October-21
Mujeres
A comentar!
Eres importante para nosotros.
Por favor ayúdanos a mejorar nuestro restaurante.
Para hacerlo solo tienes que opinar acerca de nuestros servicios
Añade un comentario
Debe ingresar mínimo 100 letras
Agregar comentario
Eres importante para nosotros.
Estamos guardando tu comentario, espera un momento por favor
acerca de nuestros servicios
Todo muy bien, en concreto el buen trato y que hiciera buen tiempo y me pusiera moreno. La comida espectacular
Agregar comentario
Eres importante para nosotros.
Tú comentario ha sido guardado
Por favor, sigue agregando más comentarios
OK
Añade un comentario
Agregar comentario

COCOments
Querido cliente, esta es tu lista de comentarios.
Por favor, sigue comentando
Eligio Sommelier
7/22/22, 12:35 PM
BUENO
Comimos un menu degustacion de seis pasos donde los vinos los eligio el sommelier
Muy Rica
7/22/22, 12:34 PM
NEUTRO
Esta muy rica pero las cantidades no son muy grandes, te quedas con


COCOments
Querido Juan, puedes filtrar tus comentaios por:
Todos
Fechas
Sentimiento
Ambos
Comentario
7/12/22, 12:52 AM
NEUTRO
Prueba de comentario dentro de Post
Prueba Comentario
7/12/22, 1:14 AM
EXCELENTE
Prueba comentario
Reportes
Comentarios
Resturante
Redes Sociales

Administrador

En este apartado se presenta las capturas del despliegue de la aplicación del módulo administrador, primero tenemos la función de reportes, segundo la función para listar todos los comentarios de la aplicación, tercero la función para obtener la información sobre los comentarios del restaurante y finalmente la adquisición de comentarios de redes sociales.



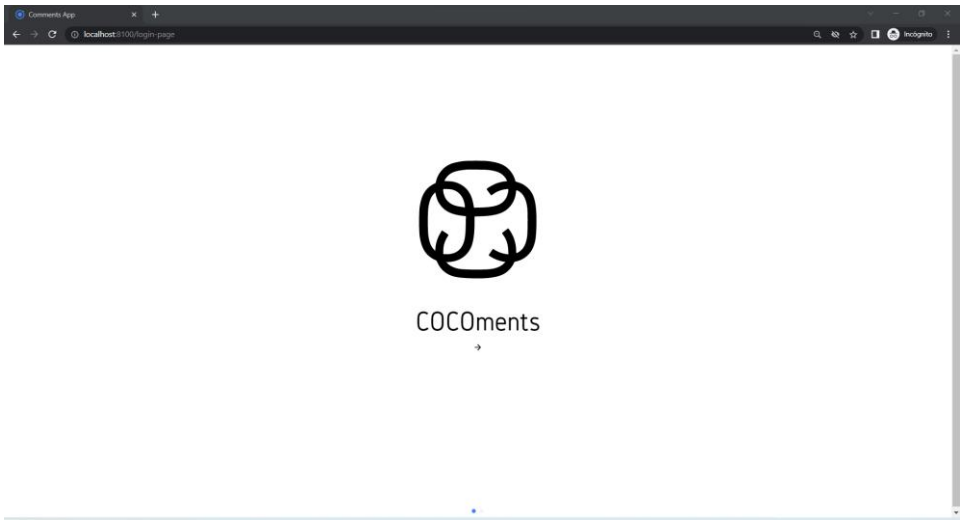
COCOments
Querido Juan,puedes analizar Tus Redes Sociales
Desea buscar por Fecha:
D... 2022-07-01 21:16:00
H... 2022-07-21 21:16:00
Buscar
Ver todos
¿Desea analizar sus nuevos comentarios de Facebook?
Analizar Ahora
Lista de Comentarios
Juan Carlos
7/14/22, 10:43 AM
Prueba commentarios con facebook
Reportes
Comentarios
Resturante
Redes Sociales
COCOments
¿Desea analizar sus nuevos comentarios de Facebook?
Analizar Ahora
Lista de Comentarios
Juan Carlos
7/14/22, 10:43 AM
Prueba commentarios con facebook
Invisible Dreamer
7/13/22, 11:29 PM
Prueba de comentario dentro de Post2
Invisible Dreamer
7/12/22, 12:52 AM
Prueba de comentario dentro de Post
Juan Carlos
7/14/22, 11:55 AM
Prueba comenarios desde Juan
Facebook Profile Private
7/12/22, 1:14 AM
Prueba comentario
Reportes
Comentarios
Resturante
Redes Sociales
Bienvenido Juan
TERRACE
Juan, puedes ver lo que opinan sus clientes con respecto a:
Seleccione Categoria
Filtrar por:
Obtener imagen
amigo precio probar este
servicio cosa
buen jecar si
vivierar
cuenco comida
delicios
propietario módico
esposo
generoso querer porción
visitar almorzar menú chef
lugar volver
excelente
amable
seguro
Buscar por palabra:
Buscar algo en específico
Reportes
Comentarios
Resturante
Redes Sociales

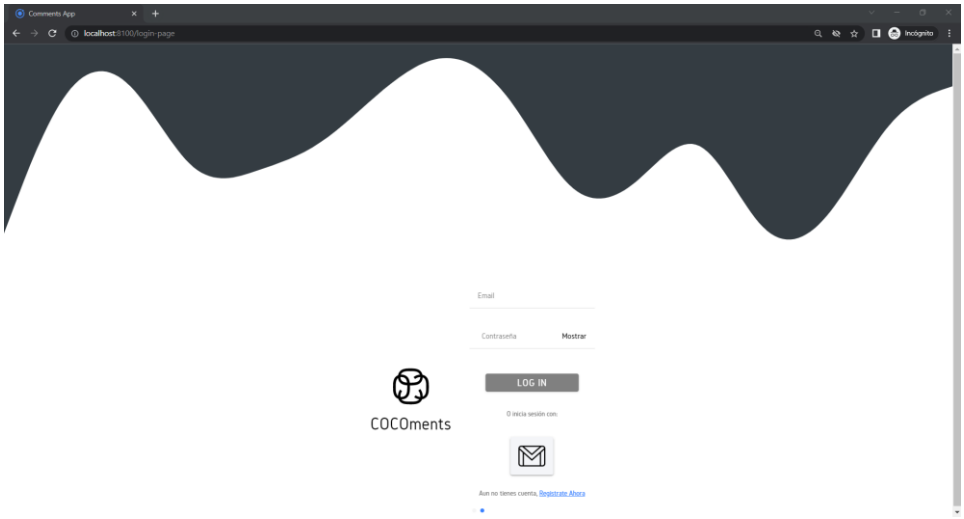
B. Aplicación Web

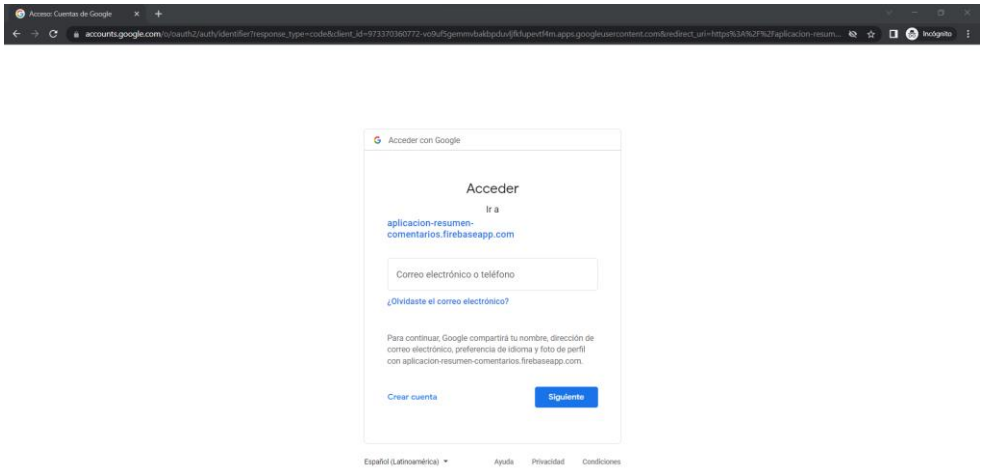
En este anexo se presenta capturas sobre el despliegue de la aplicación del módulo usuario y el módulo administrador sobre web.

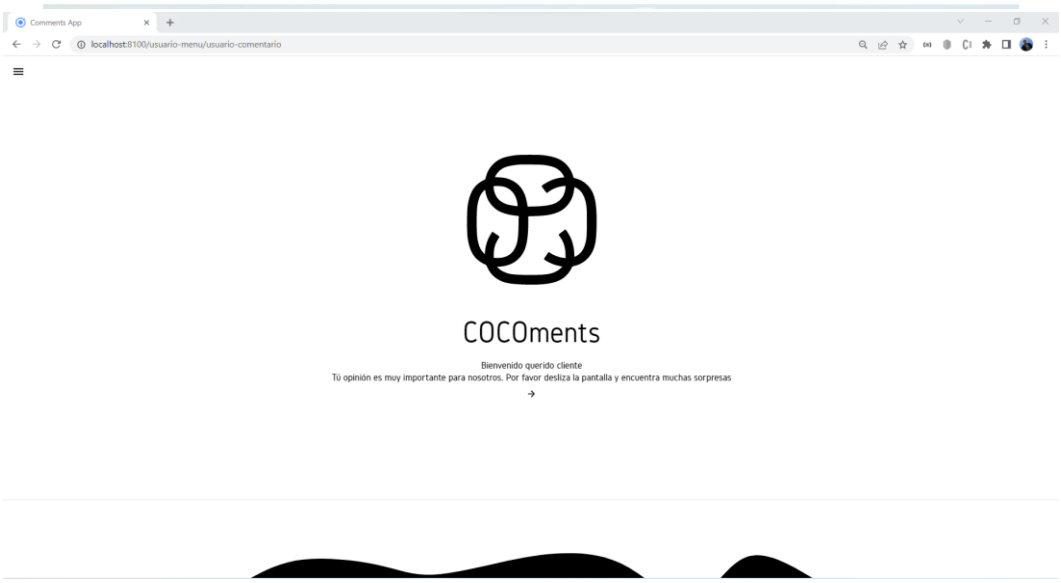
Usuario

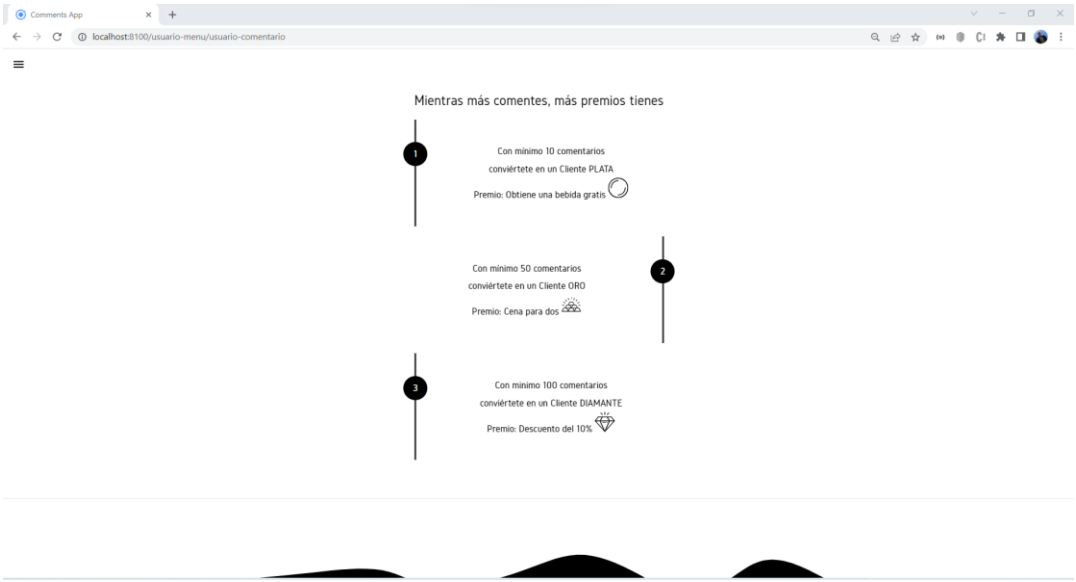
En este apartado se presenta las capturas del despliegue de la aplicación del módulo usuario, primero tenemos la función de autentificación, segundo la función registrar, tercero la función realizar comentario y finalmente un listar comentario.

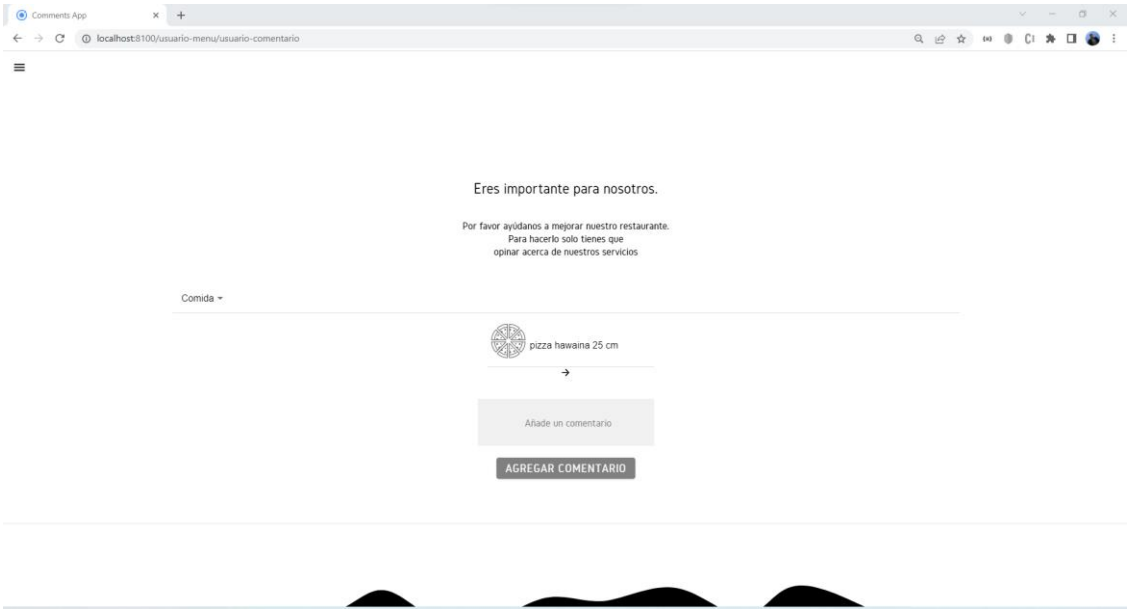


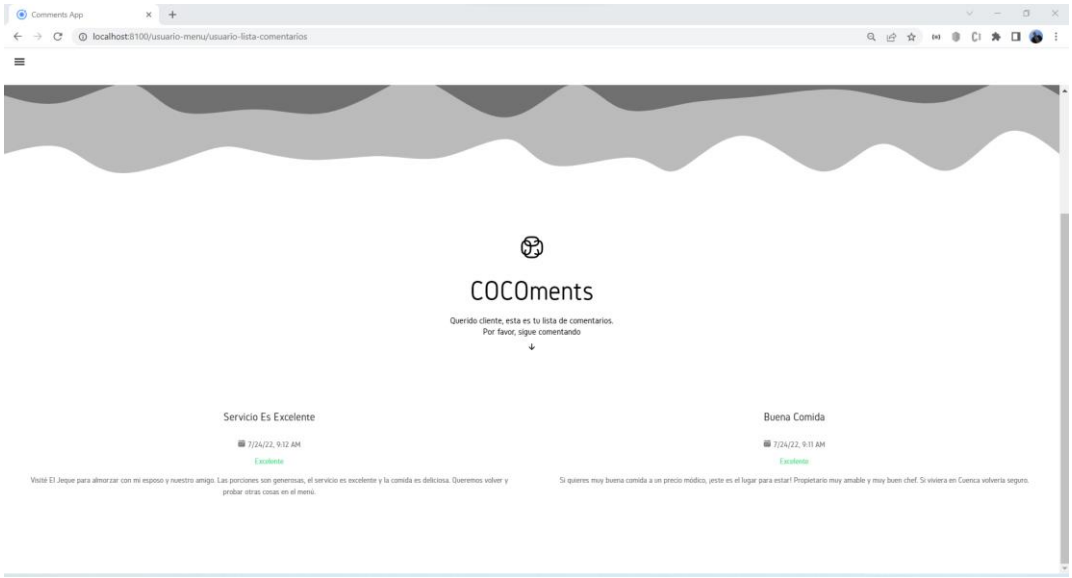


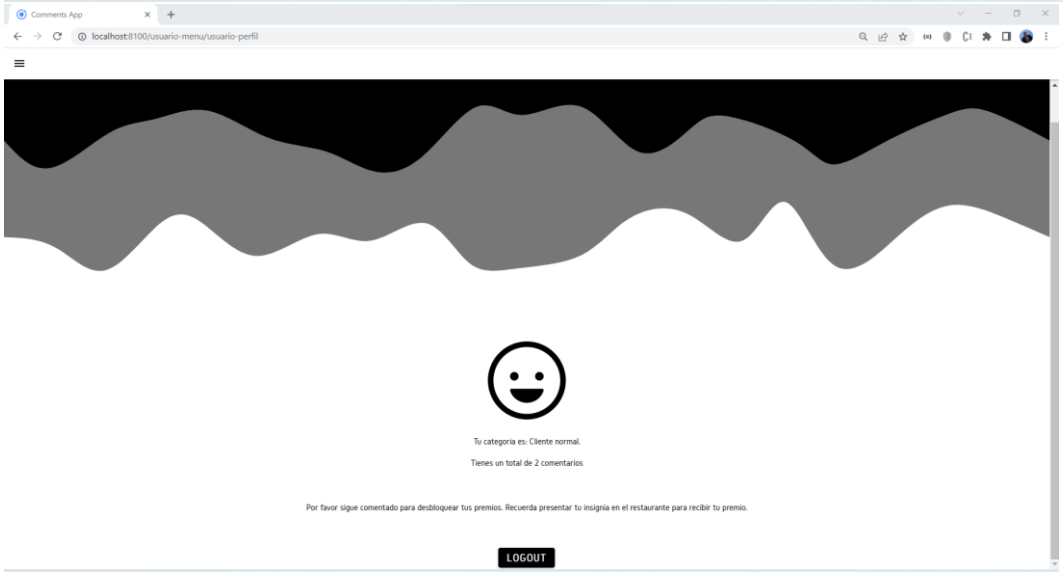

Acceso: Cuentas de Google
accounts.google.com
Incógnito
Acceder con Google
Acceder
Ir a
aplicacion-resumen-comentarios.firebaseapp.com
Correo electrónico o teléfono
¿Olvidaste el correo electrónico?
Para continuar, Google compartirá tu nombre, dirección de correo electrónico, preferencia de idioma y foto de perfil con aplicacion-resumen-comentarios.firebaseapp.com.
Crear cuenta
Siguiente
Español (Latinoamérica)
Ayuda
Privacidad
Condiciones

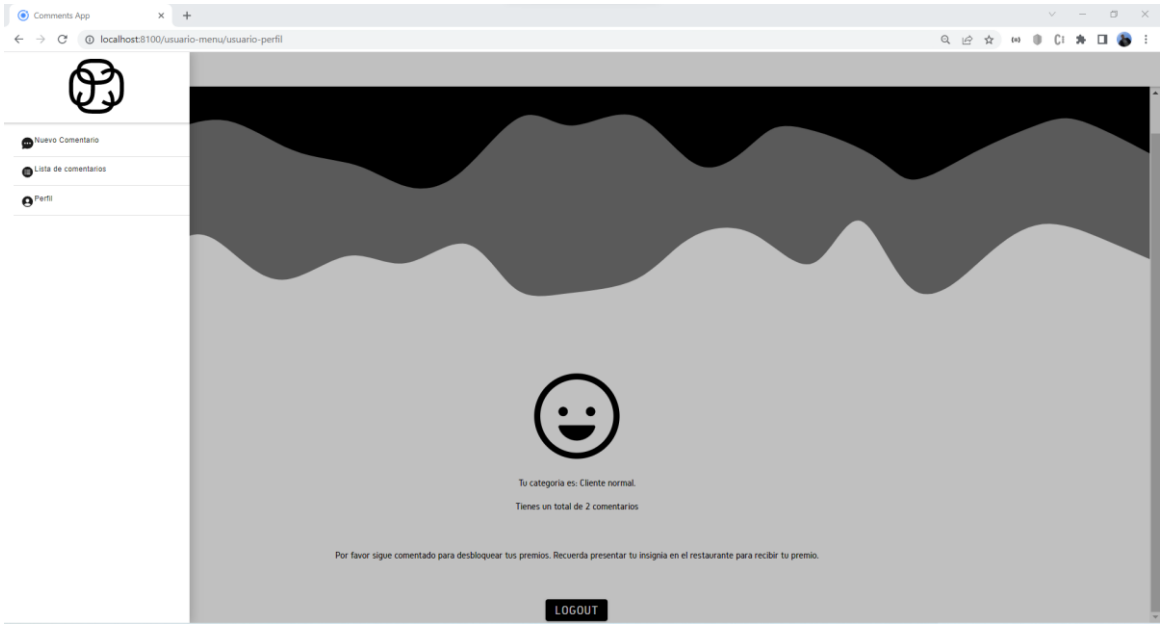

Comments App
localhost:8100/usuario-menu/usuario-comentario
COCOments
Bienvenido querido cliente
Tú opinión es muy importante para nosotros. Por favor desliza la pantalla y encuentra muchas sorpresas

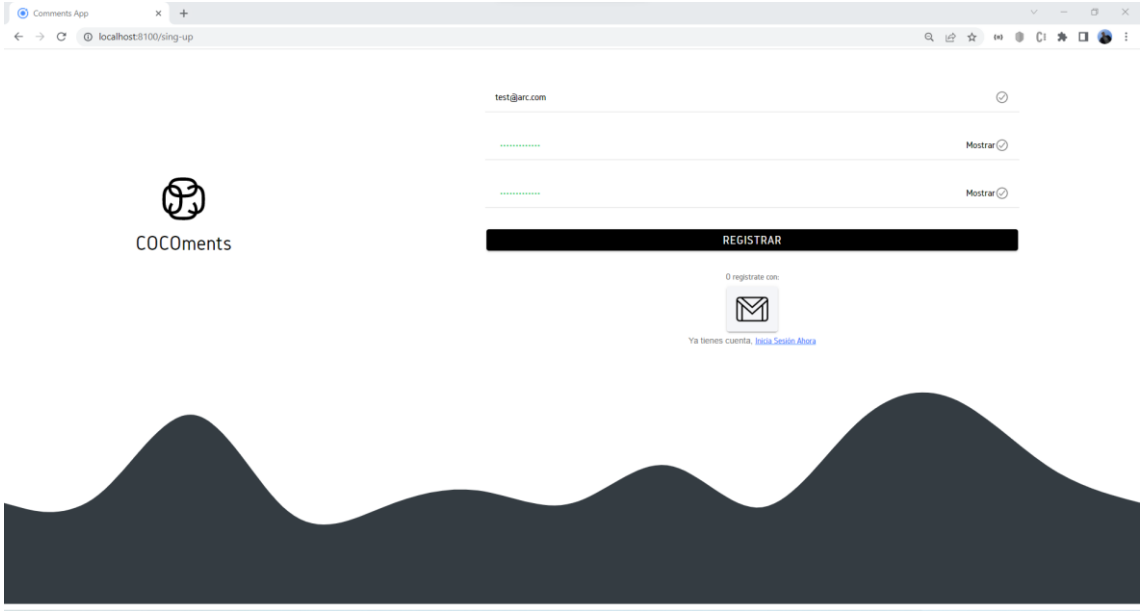

Comments App
localhost:8100/usuario-menu/usuario-comentario
Mientras más comentes, más premios tienes
1
Con mínimo 10 comentarios
conviértete en un Cliente PLATA
Premio: Obtiene una bebida gratis
2
Con mínimo 50 comentarios
conviértete en un Cliente ORO
Premio: Cena para dos
3
Con mínimo 100 comentarios
conviértete en un Cliente DIAMANTE
Premio: Descuento del 10%

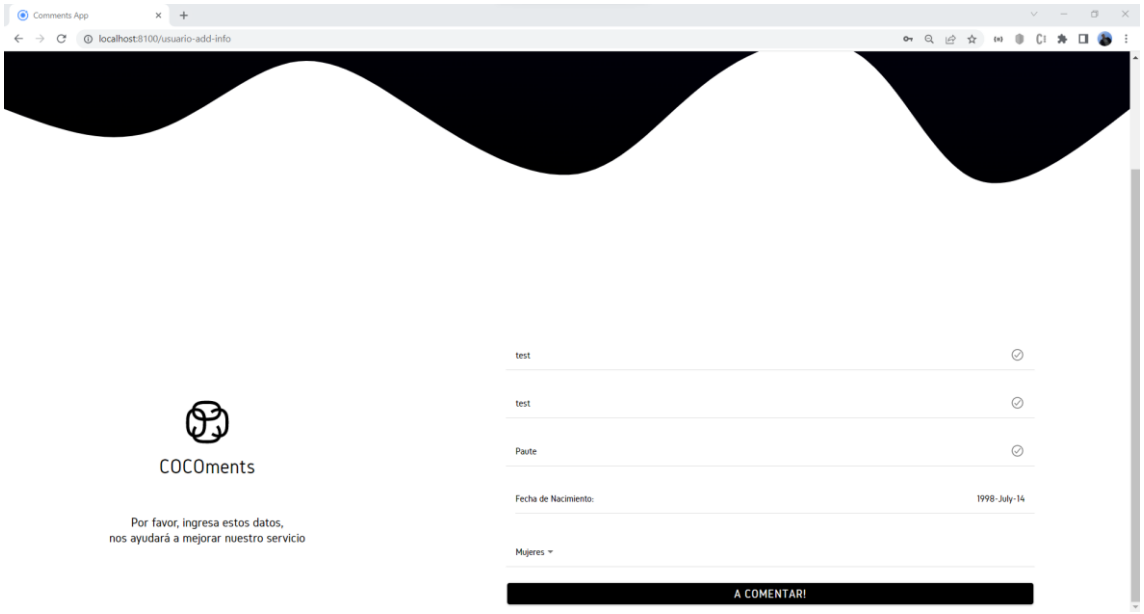

Comments App
localhost:8100/usuario-menu/usuario-comentario
Eres importante para nosotros.
Por favor ayúdanos a mejorar nuestro restaurante.
Para hacerlo solo tienes que
opinar acerca de nuestros servicios
Comida
pizza hawaina 25 cm
Añade un comentario
AGREGAR COMENTARIO


Comments App
localhost:8100/usuario-menu/usuario-lista-comentarios
COCOments
Querido cliente, esta es tu lista de comentarios.
Por favor, sigue comentando
Servicio Es Excelente
7/24/22, 9:12 AM
Excelente
Visité El Jeque para almorzar con mi esposo y nuestro amigo. Las porciones son generosas, el servicio es excelente y la comida es deliciosa. Queremos volver y probar otras cosas en el menú.
Buena Comida
7/24/22, 9:11 AM
Excelente
Si quieres muy buena comida a un precio módico, ¡este es el lugar para estar! Propietario muy amable y muy buen chef. Si viviera en Cuenca volvería seguro.


Comments App
localhost:8100/usuario-menu/usuario-perfil
Tu categoría es: Cliente normal.
Tienes un total de 2 comentarios
Por favor sigue comentado para desbloquear tus premios. Recuerda presentar tu insignia en el restaurante para recibir tu premio.
LOGOUT

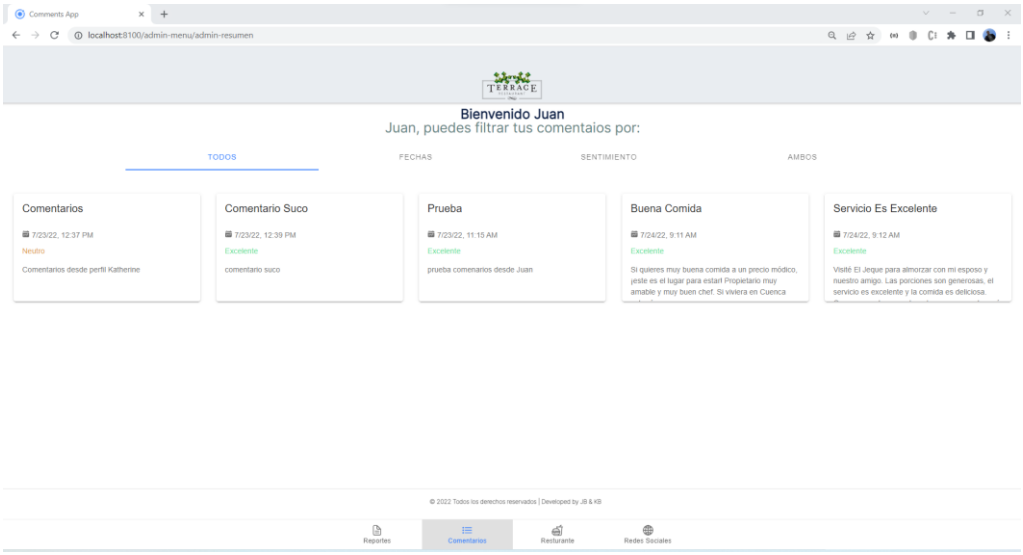

Comments App
localhost:8100/usuario-menu/usuario-perfil
Nuevo Comentario
Lista de comentarios
Perfil
Tu categoria es: Cliente normal.
Tienes un total de 2 comentarios
Por favor sigue comentado para desbloquear tus premios. Recuerda presentar tu insignia en el restaurante para recibir tu premio.
LOGOUT

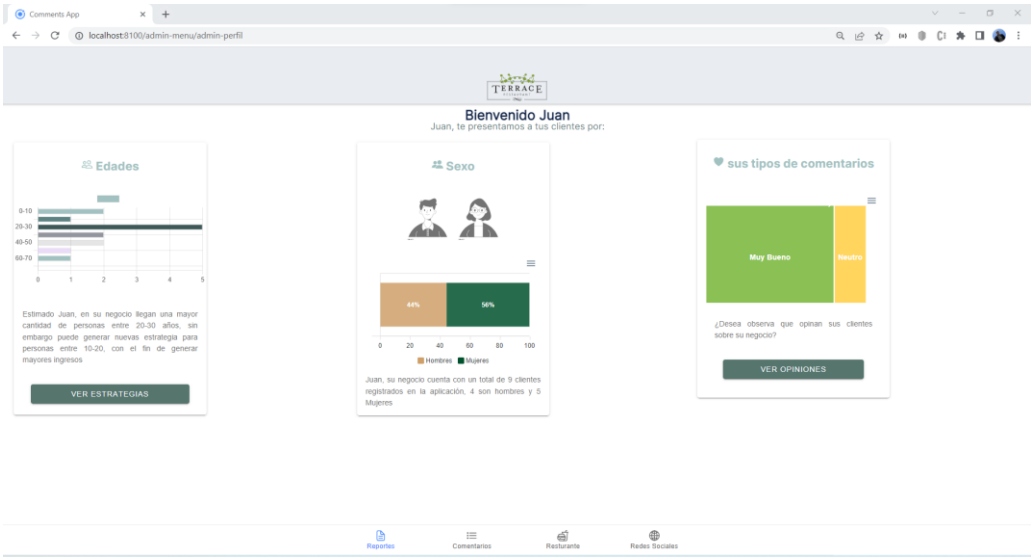

Comments App
localhost:8100/sing-up
test@arc.com
Mostrar
Mostrar
COCOments
REGISTRAR
O registrate con:
Ya tienes cuenta, Inicia Sesión Ahora


Comments App
localhost:8100/usuario-add-info
COCOments
Por favor, ingresa estos datos,
nos ayudará a mejorar nuestro servicio
test
test
Paute
Fecha de Nacimiento:
1998-July-14
Mujeres
A COMENTAR!

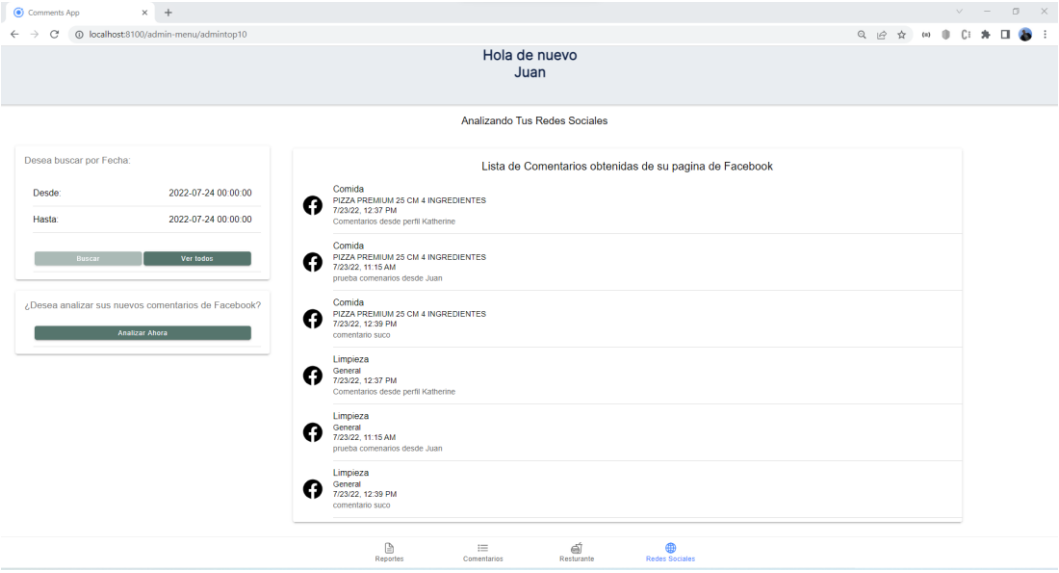
Administrador

En este apartado se presenta las capturas del despliegue de la aplicación del módulo administrador, primero tenemos la función de reportes, segundo la función para listar todos los comentarios de la aplicación, tercero la función para obtener la información sobre los comentarios del restaurante y finalmente la adquisición de comentarios de redes sociales.



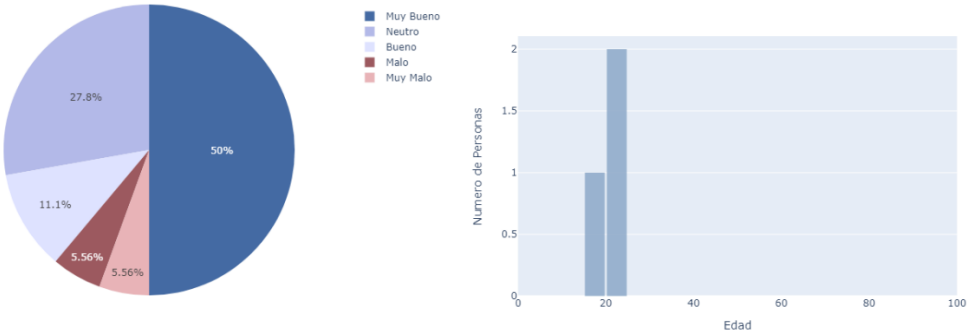


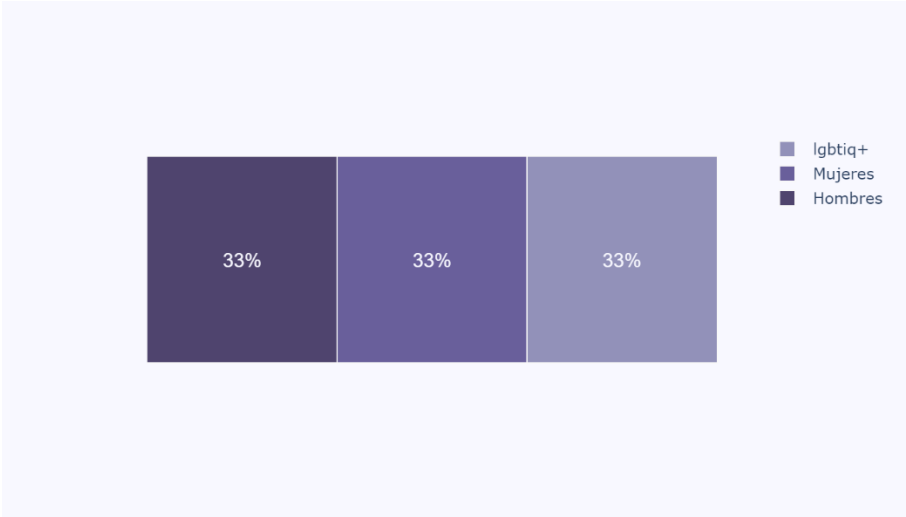
Comments App
localhost:8100/admin-menu/adminkeywords
TERRACE
Bienvenido Juan
Juan, estos son los mejores y peores platos del restaurantes:
Los Mejores Platos
Los Peores Platos
Juan, puedes ver lo que opinan sus clientes con respecto a:
Seleccione Categoria
Filtrar por:
OBTENER IMAGEN
amigo
precio
probar
este
servicio
cosa
buen
jecar
si
excelente
amable
seguro
vivierar
cuenco
comida
deliciós
Reportes
Comentarios
Resturante
Redes Sociales

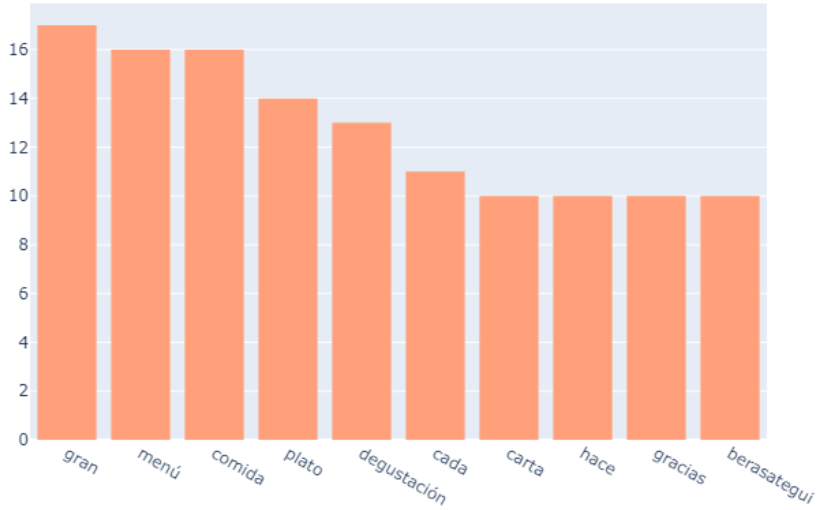

Comments App
localhost:8100/admin-menu/admintop10
Hola de nuevo
Juan
Analizando Tus Redes Sociales
Desea buscar por Fecha:
Desde: 2022-07-24 00:00:00
Hasta: 2022-07-24 00:00:00
Buscar
Ver todos
¿Desea analizar sus nuevos comentarios de Facebook?
Analizar Ahora
Lista de Comentarios obtenidas de su pagina de Facebook
Comida
PIZZA PREMIUM 25 CM 4 INGREDIENTES
7/23/22, 12:37 PM
Comentarios desde perfil Katherine
Comida
PIZZA PREMIUM 25 CM 4 INGREDIENTES
7/23/22, 11:15 AM
prueba comenarios desde Juan
Comida
PIZZA PREMIUM 25 CM 4 INGREDIENTES
7/23/22, 12:39 PM
comentario suco
Limpieza
General
7/23/22, 12:37 PM
Comentarios desde perfil Katherine
Limpieza
General
7/23/22, 11:15 AM
prueba comenarios desde Juan
Limpieza
General
7/23/22, 12:39 PM
comentario suco
Reportes
Comentarios
Resturante
Redes Sociales

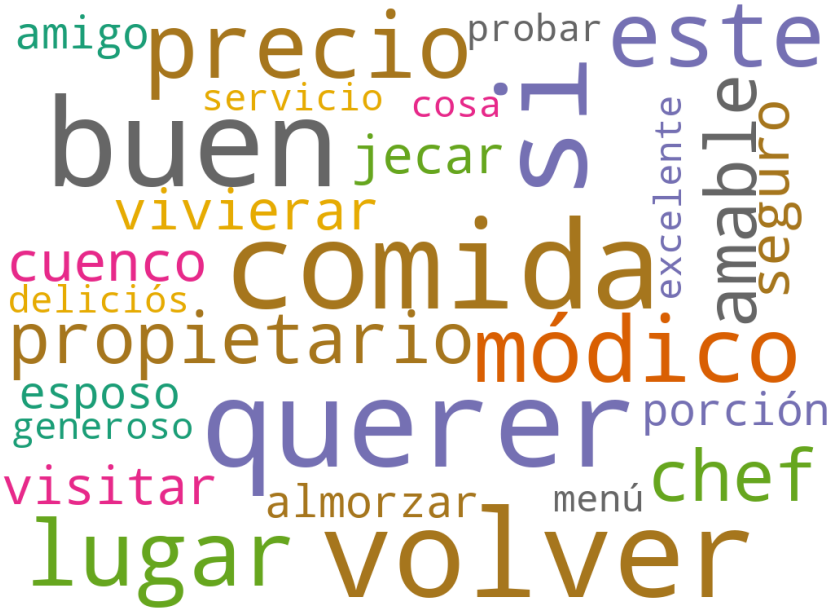
C. Gráficas generadas desde Django para análisis estadístico

En este anexo, se presenta gráficas que explican con mayor detalle cómo se encuentra el dataset que contiene a los comentarios y usuarios e información del restaurante.






16
14
12
10
8
6
4
2
0
gran
menú
comida
plato
degustación
cada
carta
hace
gracias
berasategui


amigo
precio
probar
este
servicio
cosa
buen
jecar
si
excelente
amable
seguro
vivierar
cuenco
comida
deliciós
propietario
módico
esposo
generoso
querer
porción
visitar
almorzar
menú
chef
lugar
volver

D. Codificación

Dentro de cada enlace encontrará el código del proyecto tanto para el modelo de resúmenes de texto y fronEnd - backend de la aplicación

Código de algoritmo para Resúmenes de texto

Enlace de repositorio: https://github.com/Juancarlos56/TESIS-ALGORITMO-TRANSFORMER-APLICACION_HIBRIDA_PARA_RESUMIR_Y_REPRESENTAR_COMENTARIOS.git

Código del FrontEnd para aplicación híbrida inteligente

Enlace de repositorio: https://github.com/Juancarlos56/TESIS-FRONTEND-APLICACION_HIBRIDA_PARA_RESUMIR_Y_REPRESENTAR_COMENTARIOS.git

Código del BackEnd para aplicación híbrida inteligente

Enlace de repositorio: https://github.com/Juancarlos56/TESIS-BACKEND-APLICACION_HIBRIDA_PARA_RESUMIR_Y_REPRESENTAR_COMENTARIOS.git