

**Universidad Internacional de La Rioja
Máster universitario en Ingeniería de Software y
Sistemas Informáticos**

Automatización de tareas de validación por medio de procesamiento de lenguaje natural

Trabajo Fin de Máster

Tipo de trabajo: Tipo 2. Planificación de un proyecto de desarrollo de software

Presentado por: Alvarado González, Edgar Emiro

Director/a: Sicilia Montalvo, Juan Antonio

Resumen

El presente trabajo contiene la planificación de un proyecto para automatizar las tareas de validación de análisis de sentimiento en comunicaciones de servicio con usuarios. Se analizan las herramientas de lenguaje natural disponibles actualmente para resolver este tipo de problemas, encontrando en el *Workshop SemEval* una buena referencia de modelos con referencias de comparación de resultados obtenidos. Posteriormente se analizan las alternativas a nivel de *Software* y *Hardware* para la implementación, encontrando que es conveniente usar un enfoque de uso de servicios de computación en la nube. Posteriormente, se plantean los elementos de planificación con base en el nuevo proceso sugerido para el uso de la herramienta con el objetivo de reconocer módulos, fases, historias de usuario, actividades para la planificación y elementos de diseño. Para la planificación se identifica el equipo de trabajo requerido, cronograma, presupuesto, riesgos y lecciones aprendidas. A continuación se adicionan los elementos de diseño de la solución con diagramas esquemáticos *UML*, arquitectura de la solución, interfaces de usuario, técnicas, herramientas, lenguajes de implementación, tareas para sintonía fina del modelo a usar, pruebas a realizar y validación de usabilidad. Se concluye que se cumplen los objetivos de validación de modelo, revisión de herramientas en proveedores de servicios en la nube, y planificación del flujo de trabajo. Se concluye también la importancia de establecer el proceso organizacional a implementar con base en los nuevos roles que desempeñarán tareas más analíticas que operativas para dar una perspectiva completa de los contactos de los clientes con la organización, y que la planificación puede ser flexible para la adaptación a tipos de negocio particulares.

Palabras Clave: Procesamiento de lenguaje natural, *SemEval*, modelo, computación en la nube, análisis de sentimiento.

Abstract

This work contains the planning of a project to automate the sentiment analysis validation tasks in service communications with users. The natural language tools currently available to solve this type of problem are analyzed, finding in the SemEval Workshop a good reference of models with references to compare the results obtained. Subsequently, the alternatives at the Software and Hardware level for implementation are analyzed, finding that it is convenient to use an approach of using cloud computing services. Subsequently, the planning elements are raised based on the new process suggested for the use of the tool with the aim of recognizing modules, phases, user stories and activities for planning, and design elements. For planning, the required work team, schedule, budget, risks and lessons learned are identified. Next, the design elements of the solution are added with UML schematic diagrams, solution architecture, user interfaces, techniques, tools, implementation languages, tasks for fine tuning the model to use, tests to be carried out and usability validation. It is concluded that the objectives of model validation, review of tools in cloud service providers, and workflow planning are met. The importance of establishing the organizational process to be implemented based on the new roles that will perform more analytical than operational tasks is also concluded to give a complete perspective of the customer's contacts with the organization, and that planning can be flexible for adaptation. to particular types of business.

Keywords: Natural language processing, SemEval, model, cloud computing, sentiment analysis.

Índice de contenidos

1. Introducción.....	11
1.1. Justificación	12
1.2. Planteamiento del problema	13
1.3. Estructura de la memoria	14
2. Contexto y estado del arte.....	16
2.1. Contexto del problema	16
2.2 Conjunto de soluciones complementarias.....	24
2.3 Conjunto de soluciones similares o parciales.....	27
2.4. Conclusiones	31
3. Objetivos concretos y metodología de trabajo	32
3.1. Objetivo general.....	32
3.2. Objetivos específicos	32
3.3. Metodología del trabajo	32
4. Desarrollo específico de la contribución	34
4.1. Estructura de desglose de trabajo.....	34
4.2. Estudio del dominio.....	35
4.3. Análisis y modelado de requisitos de <i>software</i>	37
4.3.1. Proceso sugerido	37
4.3.2. Proceso de consolidación de análisis.....	37
4.3.3. Proceso de validación de información.....	38
4.3.4. Proceso de gestión de información analizada	39
4.3.5. Proceso de sintonía fina de modelo	39
4.3.6. Proceso de análisis ejecutivo	40
4.4. Planificación.....	41
4.4.1. <i>Visual Story Mapping</i>	41

4.4.2. Distribución de esfuerzos.....	44
4.4.3. Calendario de ejecución.....	45
4.4.4. Presupuesto.....	49
4.4.5. Análisis de riesgo.....	51
4.4.6. Lecciones aprendidas	51
4.5. Modelado del diseño de <i>software</i>	52
4.5.1. Casos de uso.....	52
4.5.2. Requisitos funcionales y no funcionales.....	53
4.5.3. Diagrama de actividad de validación, gestión y sintonía fina.....	54
4.5.4. Diagrama de estados de <i>tickets</i> de gestión.....	56
4.5.5. Diagrama de colaboración de procesamiento de datos.....	56
4.5.6. Diagrama de componentes de sistema	57
4.6. Modelo de proceso de desarrollo	58
4.7. Arquitectura de solución	58
4.8. Interfaz de usuario	59
4.9. Técnicas, lenguajes y herramientas de implementación	63
4.10. Tareas de sintonía fina de modelo	64
4.10.1 Definición de Requisitos.....	64
4.10.2 Análisis y diseño	65
4.11. Pruebas y validación.....	65
4.11.1. Diseño de pruebas.....	65
4.11.2. Validación de usabilidad	72
5. Conclusiones y trabajo futuro	73
5.1. Conclusiones	73
5.2. Líneas de trabajo futuro	74
6. Bibliografía	75
Anexos	80

Anexo I. Artículo	80
-------------------------	----

Índice de tablas

Tabla 1. Análisis común de sentimiento basado en aspectos (ABSA) en la tarea 4 de <i>SemEval</i> 2014 con información de dominio de restaurantes.	17
Tabla 2. Análisis común de sentimiento basado en aspectos (ABSA) y análisis de sentimiento dependiente del objetivo en la subtarea 1 (<i>slot</i> 2) tarea 5 de <i>SemEval</i> 2016 con información de dominio de restaurantes.	18
Tabla 3. Calendario de ejecución	45
Tabla 4. Esfuerzo por rol	49
Tabla 5. Requisitos funcionales y no funcionales.	53
Tabla 6. Tipos de desarrollos, herramientas y lenguajes de programación	64
Tabla 7. Casos de prueba	66

Índice de figuras

Figura 1. Datos de evaluación de tarea 4 de SemEval 2014 (Manandhar, et al., 2014a).....	18
Figura 2. Resultados de uno de los participantes en la tarea 9 de SemEval 2020 (Neto, Amaral, Silva, & Soares, 2020).....	19
Figura 3. Resultados de tarea 9 de SemEval 2020 para el análisis de sentimiento de tweets en “Spanglish” (Patwa, et al., 2020)	20
Figura 4. Ejemplo de datos de tarea 9 de SemEval 2020 para el análisis de sentimiento de tweets en “Spanglish” (Patwa, et al., 2020)	20
Figura 5. Alternativas de HW para soluciones de ML y DL (Jawandhiya, 2018).	22
Figura 6. Elementos de entrada para el proyecto propuesto (elaboración propia).	23
Figura 7. Hardware propuesto por Google Cloud Platform para modelos de ML y DL – Cloud TPU v3 (Sato, 2018).....	24
Figura 8. Modelos de ML y DL para Cloud TPUs (Sato, 2018).	24
Figura 9. Solución de análisis de sentimiento propuesta por Vonage (Vonage, 2020).....	28
Figura 10. Request y response de API de sentimiento de Cloudmersive (Cloudmersive, 2020).....	29
Figura 11. Herramientas de análisis de sentimiento de Meaning Cloud (MeaningCloud, 2020)	29
Figura 12. Identificación de análisis de sentimiento basado en aspectos de la API de Meaning Cloud (Matarranz, 2017)	30
Figura 13. Identificación de análisis de sentimiento basado en aspectos de la API de análisis de sentimiento de Microsoft (Steen, et al., 2020).....	30
Figura 14. Identificación de análisis de sentimiento basado en aspectos de la API de Repustate (Repustate, 2020)	31
Figura 15. Técnica Agile v2 propuesta por framework Power Apps (Microsoft, 2020c).....	33
Figura 16. Estructura de desglose de trabajo de proyecto (elaboración propia).	35
Figura 17. Ejemplo de aspectos de análisis de sentimiento para la industria bancaria (Lappemana, Clark, Evans, Sierra-Rubiac, & Gordon, 2020)	36
Figura 18. Proceso de consolidación de análisis basado en aspectos (Elaboración propia).	38

Figura 19. Proceso de analista de validación y gestión de información (Elaboración propia)	38
Figura 20. Proceso de sintonía fina de modelo inicial (Google, 2020c)	39
Figura 21. Proceso de sintonía fina frecuente de modelo (Google, 2020c)	40
Figura 22. Visual Story Mapping – Fase de Sintonía fina inicial de modelo – Levantamiento de requisitos (Elaboración propia)	41
Figura 23. Visual Story Mapping – Fase de Sintonía fina inicial de modelo – Análisis y diseño (Elaboración propia)	42
Figura 24. Visual Story Mapping – Fase de Sintonía fina inicial de modelo – Desarrollo y pruebas (Elaboración propia)	42
Figura 25. Visual Story Mapping – Modulo de sistema de procesamiento de información de grabaciones (Elaboración propia)	43
Figura 26. Visual Story Mapping – Modulo de control de analíticas de validación (Elaboración propia)	43
Figura 27. Visual Story Mapping – Módulos de Workflow de gestión, control de roles y Fases de Pruebas integrales y Validación de usabilidad (Elaboración propia)	44
Figura 28. Calculadora de precios de Google Cloud Platform para la API de Speech to Text (Google, 2020a)	49
Figura 29. Calculadora de precios de AI Platform que permite la publicación del modelo de análisis de sentimiento basado en aspectos (Google, 2020a)	50
Figura 30. Calculadora de precios de Azure para máquina virtual con licencia de Windows (Microsoft, 2020b)	50
Figura 31. Casos de uso (Elaboración propia)	52
Figura 32. Diagrama de actividades de proceso de validación, gestión y sintonía fina frecuente de modelo (Elaboración propia)	55
Figura 33. Diagrama de estados para tickets de gestión de análisis (Elaboración propia)	56
Figura 34. Diagrama de colaboración de procesamiento de información (Elaboración propia)	56
Figura 35. Diagrama de componentes de presentación, lógica de negocio y backend (Elaboración propia)	57
Figura 36. Diagrama de arquitectura (Elaboración propia)	58

Figura 37. Interfaz inicial de proceso de consolidación de información (Elaboración propia)	59
Figura 38. Interfaz proceso de consolidación de información de registro de casos (Elaboración propia)	60
Figura 39. Interfaz proceso de consolidación de información con opciones de acciones sobre registros detallados (Elaboración propia)	60
Figura 40. Interfaz inicial de proceso de validación de información (Elaboración propia)	61
Figura 41. Interfaz que muestra los registros detallados de los casos (Elaboración propia) .	62
Figura 42. Interfaz de reproducción de grabación (Elaboración propia)	62
Figura 43. Burndown Chart de JIRA (Rehkopf, 2020)	63
Figura 44. Encuesta System Usability Scale (SUS) (Brooke, 2020)	72

1. Introducción

La gran mayoría de organizaciones atienden usuarios. Es importante la conexión de los altos ejecutivos con la opinión de los usuarios, para entender como se está desempeñando la cadena de valor de la organización para solucionar las necesidades de los usuarios en procesos como venta y posventa.

NLP (Procesamiento de lenguaje natural) permite a las empresas recopilar información y mejorar su servicio. Alrededor del 40% de las grandes empresas lo han adoptado o lo adoptarán a finales de 2019. (Newman, 2018)

Las técnicas de procesamiento de lenguaje natural permiten alcanzar objetivos de automatización y análisis en las comunicaciones a usuarios que permiten tener una perspectiva global de las comunicaciones modelando lo que es importante para una organización en particular.

Actualmente las empresas hacen esfuerzos orientados en implementar ChatBots, hacer análisis predictivos, usar los análisis para controlar variables como *Churn* y Fraude. En este trabajo se propone hacer uso de modelos de NLP para sintetizar en perspectivas de análisis manejables la comunicación de usuarios, con esto se quiere obtener la mayor transparencia posible en el análisis. En la academia y los fabricantes líderes en tecnología se tienen propuestas para solucionar problemas como este. Para alcanzar estos objetivos es importante que las organizaciones estén dispuestas a cambiar sus procesos de atención para cambiar las prioridades de procesos operativos a procesos analíticos para dar un alcance mayor a las perspectivas.

Para la planificación del proyecto es importante el manejo de técnicas de ingeniería de software para documentar los requisitos y evitar ambigüedades en el entendimiento de estos. También es importante el uso de técnicas ágiles para gestionar requisitos que pueden ser cambiantes, pero al mismo tiempo se planifica usar técnicas de levantamiento y una fuerte base conceptual para dar al proyecto una buena base teórica y de conocimiento de negocio.

1.1. Justificación

El objetivo del presente trabajo es planificar un proyecto para transformar el trabajo de analistas que validan información de sentimiento de comunicaciones de clientes / usuarios, en un trabajo más analítico y menos operativo por medio del uso de *APIs* (Interfaces de programación de aplicaciones) y modelos pre-existentes de procesamiento de lenguaje natural. Se pretende solventar el problema de usar recursos humanos para trabajos repetitivos que pueden ser automatizados en un buen porcentaje.

Existen modelos pre-existentes y *APIs* que analizan información y entregan una salida que es necesario procesar dependiendo del contexto para encajar en un entorno corporativo. La novedad es usar las herramientas novedosas existentes para planificar el proyecto para su implementación, complementándolas con herramientas convencionales que ya usan la mayoría de los entornos corporativos. Por ejemplo, combinar una *API* de lenguaje natural (novedosa) con un flujo de trabajo con roles y control de acceso por usuario (convencional). Por ejemplo, combinar el resultado de la dispersión de la clasificación de modelos de lenguaje natural (novedoso) con los indicadores de cumplimiento y nivel de servicio de respuestas (convencional).

En organizaciones que trabajan productos y servicios “*commodities*” es más importante tener un diferencial de servicio, porque hay competidores que venden prácticamente el mismo producto. Para obtener un diferencial en experiencia es importante entender la opinión de los usuarios acerca de cada uno de los aspectos de los productos y servicios. Las organizaciones actualmente tienen forma de evaluar algunos aspectos con base en indicadores convencionales, como abandono de llamadas, tiempo de atención, tiempo de servicio, cancelación de servicio, cantidades de ventas. El objetivo del trabajo es analizar los aspectos que evalúan los usuarios de productos y servicios directamente, usando sus propias palabras.

Además del análisis propuesto, hay potencial de abrir nuevas oportunidades con el análisis de información que seguramente aprovecharán los líderes funcionales apalancados por el uso de herramientas tecnológicas, tal como el marketing experiencial (Alfaro, 2010). Si se puede medir la experiencia de los usuarios con indicadores basados en sus propias palabras, va a ser posible mejorarla.

1.2. Planteamiento del problema

Teniendo en cuenta la característica del proyecto se sugiere el uso de técnicas ágiles por la incertidumbre de los requisitos y el uso de herramientas novedosas. Una de las fases del trabajo es explorar las herramientas que puedan tener bajas barreras de implementación para realizar la planificación. Es importante que la adaptación no requiera de tareas para las cuales necesitamos a un científico de datos elaborando herramientas demasiado complejas que puedan aumentar el costo de propiedad y de mantenimiento de la solución. El proceso de adaptación de las herramientas debe estar basado en un estándar del mercado que permita el manejo por parte de ingenieros de datos expertos en herramientas líderes en el mercado, que hacen que la solución integral sea reproducible.

Es importante que el modelo a usar sea transparente en cuanto a los resultados del modelo en tareas retadoras y que el entrenamiento permita mejorar los resultados del análisis continuamente. Se propone planificar la implementación de un modelo de análisis de sentimiento basado en aspectos para validar las comunicaciones entrantes de clientes en audios, haciendo uso de una *API* de reconocimiento de voz, y un modelo de *NLP* pre-entrenado para hacer la evaluación cuantitativa del análisis de sentimiento en diferentes aspectos importantes para el negocio.

El problema se divide básicamente en las siguientes partes:

- Encontrar a nivel del estado del arte el modelo a usar de procesamiento de lenguaje natural para análisis de sentimiento basado en aspectos.
- Analizar las alternativas disponibles para planificar la solución a nivel de *Hardware* y *Software*, metodología de implementación y entrenamiento continuo.
- Formular nuevos procesos definiendo los roles para pasar de procesos operativos a procesos analíticos.
- Planificar la fase de análisis de requisitos de negocio para contextualizar la sintonía fina del modelo y el flujo de trabajo.
- Formular los casos de prueba funcionales y no funcionales y la validación de usabilidad.

1.3. Estructura de la memoria

La estructura del trabajo es la siguiente:

- Contexto y estado del arte: Se hace un análisis a nivel de estado del arte de las herramientas de *NLP* disponibles para hacer el análisis requerido en el presente trabajo, se analizan las herramientas disponibles que hacen parte de la solución, y herramientas que cumplen con el objetivo parcialmente.
- Objetivos concretos y metodología de trabajo: Se plantean los objetivos que se buscan alcanzar con el plan a desarrollar en el trabajo y la metodología.
- Desarrollo específico de la contribución

En este capítulo se describe el proceso de desarrollo de la solución para detallar la planificación de manera asertiva.

- Estructura de desglose de trabajo: Se hace un resumen de las fases y módulos a considerar en el desarrollo del proyecto.
- *Visual Story Mapping*: Se detallan las fases y módulos a nivel de historias de usuario y actividades.
- Distribución de esfuerzos: Se evalúan los refuerzos requeridos y el personal a tener en cuenta en el equipo del proyecto.
- Tareas: Teniendo en cuenta la información de los numerales anteriores, se plantean las tareas para el plan de trabajo.
- Calendario de ejecución: Se planifican las fechas de ejecución de desarrollo de la solución, se cuantifican las horas requeridas por cada rol.
- Presupuesto: Costo de licencias requeridas.
- Análisis de riesgo: Se analizan las diferentes fases en las que el proyecto pueda tener variaciones de recursos y tiempo.
- Lecciones aprendidas: Se definen los puntos clave de la planificación y desarrollo en los que podría plantear mejoras en cada *Sprint*.

- Estudio del dominio: Se complementa la información del estado del arte con el análisis de los aspectos a usar para el análisis de sentimiento.
- Análisis y modelado de requisitos de *software*: Aunque en el plan se propone una solución estándar de análisis de sentimiento basado en aspectos, se detalla el análisis de modelado de requisitos de un contexto particular. Esta información es importante para establecer la necesidad de datos de sintonía fina, y el requisito del flujo de trabajo de análisis.
- Modelado del diseño de *software*: Una vez establecida la necesidad, se establece el proceso de análisis. En el caso del plan propuesto se detalla el diseño para una solución estándar de análisis.
- Arquitectura de solución: Se usan las herramientas de ingeniería de *software* para documentar la arquitectura de la solución.
- Modelo de proceso de desarrollo: Se propone el uso de un modelo pre-entrenado, y un *framework* de desarrollo de aplicaciones, que permiten planificar un tiempo relativamente corto para la obtención de entregas de valor.
- Pruebas y validación de usabilidad: Se describen los casos de prueba en detalle y la metodología de validación de usabilidad de la herramienta.
- Conclusiones y líneas de trabajo futuro: Las conclusiones responden que se cumplieron los objetivos en la planificación. Se sugieren líneas de trabajo futuro para dar más alcance a la herramienta y usar esquemas similares para diferentes dominios.

2. Contexto y estado del arte

El presente capítulo tiene por objetivo establecer el estado del arte de soluciones de *NLP* para hacer el análisis requerido, y a su vez, explorar en los proveedores de soluciones complementarias para establecer la solución a nivel de *Hardware* y *Software*. También se hace una investigación sobre soluciones similares generales y parciales.

2.1. Contexto del problema

En la actualidad existen labores de revisión y valoración de información que se utilizan para priorizar servicios y orientar estrategias organizacionales. Puede que la labor de organizaciones de servicios se oriente a priorizar la atención de los usuarios más enojados, o que tengan una necesidad relacionada a ciertos servicios, o que mencionen ciertos temas en sus comunicaciones. Gran parte de las organizaciones usan procesos manuales de valoración para el manejo de las comunicaciones corporativas. Esta labor puede demandar gran cantidad de esfuerzo, tareas repetitivas, evaluación parcializada, y precisión cuestionable.

Por otro lado, pueden existir los siguientes tipos de organizaciones que quieren sobrevivir a la era actual de uso de información masiva,

- Organizaciones que manejan productos, servicios y proyectos convencionales, que gastan una buena parte de su presupuesto desarrollando proyectos de tecnología de información con el objetivo de organizar y automatizar procesos de negocio, y
- Organizaciones que manejan productos, servicios y proyectos de innovación como “core” de negocio.

En ambos tipos de organizaciones puede adaptarse el uso de herramientas vanguardistas de inteligencia artificial para analizar información masiva.

En la academia existen avances considerables en herramientas de inteligencia artificial para enfrentar desafíos cada vez más complejos de procesamiento de lenguaje natural. Con el avance de las capacidades computacionales se ha logrado consolidar modelos de redes neuronales que pueden compararse por medio de tareas estándar para conocer de forma cuantitativa la precisión en el procesamiento de lenguaje natural. Existe un *Workshop* internacional de Evaluación Semántica llamado *SemEval* (Manandhar, et al., 2014b). En este evento anual se plantean tareas de análisis de lenguaje por medio de herramientas de

automatización. Uno de los retos que plantea el presente trabajo es abordar el análisis de sentimiento automatizado en la validación de comunicaciones corporativas. Las siguientes tablas comparan la precisión de diferentes modelos para el análisis de sentimiento basado en aspectos de la información proveniente de las opiniones de clientes sobre restaurantes en dos tareas propuestas en el evento *SemEval* (Li, Li, Zhang, Hu, & Hu, 2019):

Tabla 1. Análisis común de sentimiento basado en aspectos (ABSA) en la tarea 4 de *SemEval* 2014 con información de dominio de restaurantes.

Año	Modelo	Idea básica	Precisión
2016	<i>AE-LSTM</i>	Las palabras objetivo dadas en cada frase del cuerpo de entrenamiento son vectorizadas y adicionadas al modelo <i>LSTM</i> (Red neuronal artificial de memoria de corto plazo) como una entrada para entrenarse juntos.	76.2%
2016	<i>AT-LTSM</i>	Un mecanismo de atención es proporcionado para capturar partes clave de una frase relacionadas a un aspecto dado.	77.9%
2018	<i>AF-LSTM</i>	Una nueva capa de asociación que define dos operadores de correlación, convolución circular y correlación cíclica, es introducida para aprender relaciones entre palabras y aspectos de frases.	75.44%
2015	<i>TD-LSTM</i>	Dos redes <i>LSTM</i> son adoptadas para modelar por separado, basadas en el contexto de palabras antes y después para tareas de análisis de sentimiento dependiente del objetivo.	75.63%
2015	<i>TC-LSTM</i>	Con base en <i>TD-LSTM</i> , la información de la palabra objetivo es adicionada como una entrada.	76.01%
2016	<i>ATAE-LSTM</i>	Con base en <i>TD-LSTM</i> , la información de aspectos es introducida en dos partes del modelo: Parte de entrada y parte oculta. ATAE es Basado en atención con Aspecto embebido.	77.2%
2017	<i>IAN</i>	Aprende la atención en las palabras del objetivo y el contexto interactivamente, y genera las representaciones para objetivos y contextos separadamente. IAN es redes de atención interactiva.	78.6%
2016	<i>Memnet(k)</i>	Utiliza una red de memoria profunda con múltiples capas computacionales (saltos) para clasificar los sentimientos a nivel de aspecto, donde k es el número de capas. El objetivo principal de la red Memnet es restauración de imágenes,	(k = 2) 78.61% (k = 3) 79.06%
2018	<i>Coatención - LSTM</i> <i>Coatención - Memnet</i> (3)	Un mecanismo colaborativo de atención es propuesto para que alternativamente se usen mecanismos de atención a nivel de objetivo y a nivel de contexto	78.8% 79.7%
2017	<i>BILTSM-ATT-G</i>	Basado en el modelo de atención vainilla, este modelo es extendido para diferenciar los contextos de izquierda y derecha, y usa el método <i>gate</i> para controlar la salida de las corrientes de datos	79.73%
2018	<i>Tnet-LF</i> <i>Tnet-AS</i>	La red neuronal de convolución (<i>CNN</i>) es usada para reemplazar la red neural recurrente basada en aspectos (<i>RNN</i>) para extraer características, y estructurar la transformación preservando el contexto (<i>CPT</i>) tal como el envío sin pérdidas (<i>LF</i>) y escala adaptativa (<i>AS</i>) es usada para capturar el objetivo de información de entidades y retener la información de contexto. Tnet es redes de transformación.	80.79% 80.69%
2018	<i>AE-DLSTMs</i>	Sobre la base de <i>AE-LSTMs</i> , este modelo captura información semántica contextual en direcciones hacia adelante y hacia atrás en palabras de aspecto.	79.57%
2018	<i>AELA-DLSTMs</i>	Basado en <i>AE-DLSTMs</i> , este modelo introduce el peso de la información de posición de contexto del aspecto de palabra.	80.35%
2018	<i>StageI + StageII</i>	Introduce un mecanismo de atención de posición basado en contexto de posición entre aspecto y contexto, y también considera la perturbación de otros aspectos en la misma frase	80.10%
2018	<i>DMN + AttGRU (k=3)</i>	Una red de memoria dinámica que usa múltiples bloques de atención de múltiples mecanismos de atención es propuesta para extraer características relacionadas con sentimiento en información de memoria, donde k son los pasos de atención	81.41%
2018	<i>MGAN</i>	Este modelo diseña una pérdida de alineación de aspecto para representar interacciones a nivel de aspecto entre aspectos con el mismo contexto, y para fortalecer las diferencias de atención entre aspectos con el mismo contexto y diferentes polaridades de sentimiento. MGAN es Red de atención multigranular.	81.25%

Fuente: (Manandhar, et al., 2014b)

Tabla 2. Análisis común de sentimiento basado en aspectos (ABSA) y análisis de sentimiento dependiente del objetivo en la subtask 1 (slot 2) tarea 5 de *SemEval* 2016 con información de dominio de restaurantes.

Año	Modelo	Idea básica	Precisión
2019	BERT	Transformador bidireccional (BERT) se amplía con una capa adicional específica de la tarea y ajustada en cada tarea final BERT es representaciones de codificadores bidireccionales de transformadores.	81.54%
2019	BERT-PT	Sobre la base de BERT, son usados dos objetivos de pre-entrenamiento: Modelo de lenguaje de enmascaramiento (MLM) y predicción de la siguiente oración (NSP), para post-entrenar el conocimiento del dominio, y otra tarea de comprensión de lectura de máquina (MRC).	84.95%

Fuente: (Pontiki, Galanis, Papageorgiou, Manandhar, & Androutsopoulos, 2016)

De las anteriores tablas de los resultados de las tareas mencionadas de *SemEval* surgen al menos las siguientes preguntas:

- Los modelos entrenados para las tareas son aplicables para todos los dominios?
- Qué metodología se usa para evaluar la precisión de los modelos?
- Los modelos mencionados son multilingüaje?

Acerca de los dominios, es importante resaltar que cada tarea de *SemEval* tiene un contexto, tal como un requisito de una necesidad. Para la tarea 4 de *SemEval* 2014 se tienen los datos de entrada en XML (Lenguaje de marcado extensible):

```

<sentences>
  <sentence id="813">
    <text>All the appetizers and salads were fabulous, the steak was mouth watering and the pasta was delicious!!!</text>
    <aspectTerms>
      <aspectTerm term="appetizers" polarity="positive" from="8" to="18"/>
      <aspectTerm term="salads" polarity="positive" from="23" to="29"/>
      <aspectTerm term="steak" polarity="positive" from="49" to="54"/>
      <aspectTerm term="pasta" polarity="positive" from="82" to="87"/>
    </aspectTerms>
    <aspectCategories>
      <aspectCategory category="food" polarity="positive"/>
    </aspectCategories>
  </sentence>
  <sentence id="1579">
    <text>And really large portions.</text>
    <aspectTerms>
      <aspectTerm term="portions" polarity="positive" from="17" to="25"/>
    </aspectTerms>
    <aspectCategories>
      <aspectCategory category="food" polarity="positive"/>
    </aspectCategories>
  </sentence>
  <sentence id="2707">
    <text>Go inside and you won't want to leave.</text>
    <aspectCategories>
      <aspectCategory category="anecdotes/miscellaneous" polarity="positive"/>
    </aspectCategories>
  </sentence>
  <sentence id="3126">
    <text>Save yourself the time and trouble and skip this one!</text>
    <aspectCategories>
      <aspectCategory category="anecdotes/miscellaneous" polarity="negative"/>
    </aspectCategories>
  </sentence>

```

Figura 1. Datos de evaluación de tarea 4 de *SemEval* 2014 (Manandhar, et al., 2014a).

El modelo funciona en idioma inglés, tal como se puede apreciar en la estructura de datos, se detalla en la etiqueta *<text>* la opinión del comensal, y en las etiquetas *<aspectTerm>* y *<aspectCategories>* se detalla cada aspecto con su valoración. Esta información es la entrada del modelo para comparar el resultado con la valoración. Se maneja un mismo formato para los datos de entrenamiento y de evaluación.

Así como en idioma inglés, hay un avance importante en el análisis de sentimiento en otros idiomas, incluso análisis en textos que incluyen varios idiomas. Existe la tarea 9 de *SemEval* 2020 que tiene como objetivo evaluar el sentimiento de tweets con idiomas mezclados, específicamente Hindi e Inglés, o *Hinglish*, y Español e Inglés, o *Spanglish*. Uno de los participantes en esta tarea (Neto, Amaral, Silva, & Soares, 2020) usó la combinación de los resultados d 4 modelos, *MultiFIT* (Sintonía fina del modelo de idioma multilingüe), *BERT*, *ALBERT* (BERT ligero) y *XLNet* (método de preentrenamiento autorregresivo generalizado que permite aprender contextos bidireccionales maximizando la probabilidad esperada sobre todas las permutaciones del orden de factorización (Yang, et al., 2019)). Este modelo alcanzó una calificación F1 de 72.7%, más adelante se detallarán las fórmulas para obtener las calificaciones. El modelo *BERT* pre-entrenado puede ser “sintonizado” adicionando una capa en la salida, lo cual puede ser usado en análisis de sentimiento y otras tareas de procesamiento de lenguaje natural. El resultado del ensamble de los modelos es mejor al resultado de cada uno:

Model	F1	P	R	Acc
Ensemble	0.727	0.729	0.726	0.723
XLNet	0.679	0.696	0.692	0.690
ALBERT	0.679	0.684	0.676	0.675
BERT	0.675	0.680	0.672	0.670
MultiFiT	0.665	0.665	0.669	0.662

Figura 2. Resultados de uno de los participantes en la tarea 9 de *SemEval* 2020 (Neto, Amaral, Silva, & Soares, 2020)

Las métricas de *precision*, *recall* y F1 para estos trabajos, y como estándar para resultados de modelos de *machine learning* se definen así (Patwa, et al., 2020):

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (1)$$

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (2)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

Los resultados de las fórmulas (1) y (2) se usan para el cálculo de F1 en la fórmula (3). La mayor calificación obtenida en esta tarea fue de 80,6% para el data set de Spanglish:

Rank	System	Positive			Neutral			Negative			Avg. F1
		P	R	F1	P	R	F1	P	R	F1	
1	LiangZhao	88.3	92.6	90.4	18.1	20.9	19.4	59.9	39.5	47.6	80.6
2	rachel	89.0	87.7	88.3	16.0	45.1	23.7	65.3	24.5	35.7	77.6
3	asking28	84.5	90.1	87.2	6.1	4.9	5.4	43.5	29.9	35.4	75.6
4	dpalominop	91.6	77.2	83.8	12.7	30.6	17.9	42.9	58.6	49.5	75.5
5	kongjun	87.1	84.6	85.9	11.1	30.1	16.2	56.1	27.2	36.6	75.3
6	HaoYu	92.9	74.0	82.4	11.9	48.1	19.1	55.2	55.0	55.1	75.2
7	Taha	84.7	89.5	87.0	51.9	20.5	29.4	10.4	17.5	13.0	75.1
8	meiyim	93.0	73.3	82.0	12.1	55.8	19.9	57.7	47.1	51.9	74.5
9	Lavinia_Ap	82.0	97.9	89.2	13.8	3.9	6.1	56.0	8.0	14.1	74.4
10	jupitter	93.6	71.8	81.3	11.0	53.9	18.2	58.1	47.9	52.5	73.9
11	tangmen	91.8	72.5	81.0	11.3	55.3	18.8	59.8	41.6	49.0	73.0
12	hermosillo748	85.4	81.3	83.3	7.3	21.8	10.9	54.1	26.6	35.7	72.8
13	harsh_6	87.7	77.9	82.5	9.5	23.3	13.5	36.1	39.1	37.5	72.5
14	francesita	80.9	99.5	89.2	8.7	1.0	1.7	0.0	0.0	0.0	72.2
15	ajason08	90.1	71.0	79.4	8.2	40.3	13.6	54.7	37.5	44.5	71.0
23	Baseline	89.5	63.0	74.0	7.9	49.5	13.6	47.0	31.0	37.4	65.6

Figura 3. Resultados de tarea 9 de SemEval 2020 para el análisis de sentimiento de tweets en “Spanglish” (Patwa, et al., 2020)

Es interesante observar un ejemplo de los datos de entrenamiento para esta tarea, para observar un ejemplo del gran reto que representa la automatización del análisis de sentimiento:

```

meta      27      positive
Manager lang1
got      lang1
us       lang1
tacos   lang2
and     lang1
everyone      lang1
is       lang1
just    lang1
chilling      lang1
lol     lang1

```

Figura 4. Ejemplo de datos de tarea 9 de SemEval 2020 para el análisis de sentimiento de tweets en “Spanglish” (Patwa, et al., 2020)

En el ejemplo de datos de la figura anterior se observa que la primera columna contiene la frase, la segunda el lenguaje, y al final de la primera fila se da la calificación del sentimiento global de la frase. El objetivo de los modelos es predecir esta calificación.

Para esta tarea, así como para la mencionada anteriormente de análisis de sentimiento basado en aspectos, se encontraron los mejores resultados en modelos basados en *BERT* (Patwa, et al., 2020).

BERT es el modelo de inteligencia artificial creado por *Google* para entender el lenguaje de las búsquedas de los usuarios (CE Noticias Financieras English, 2019). *Google* liberó el código fuente para que un usuario pueda crear una variedad de modelos de análisis de texto de nivel del estado del arte (Gupta, 2018).

Para el uso de *BERT* se requiere dos tareas pre-entrenamiento y sintonía fina (Devlin, et al., 2020):

- Pre-entrenamiento: Es un proceso que requiere gran cantidad de procesamiento, pero se hace una vez por cada lenguaje. La gran mayoría de aplicaciones no requieren de una tarea de pre-entrenamiento.
- Sintonía fina: No es un proceso pesado. Una tarea convencional de *NLP* puede tomar una hora en una *Cloud TPU*, y unas pocas horas en una *GPU*. Más adelante se revisará en detalle las opciones de *hardware*.

Para el uso de *BERT* en idioma español se encuentran las siguientes opciones:

- *BERT* Multilenguaje (Devlin, et al., 2020)
- BETO (Cañete, et al., 2020a)

Teniendo en cuenta lo anterior, en los siguiente capítulos se procede a profundizar en el uso del modelo *BERT* y soluciones complementarias y similares para planificar el presente trabajo.

Una vez explorado el resultado de los modelos en tareas similares, se procede a analizar cómo se implementará teniendo en cuenta experiencias anteriores en la industria y herramientas similares.

En el presente trabajo se explorarán las alternativas sobre los siguientes aspectos de implementación:

- *Hardware*. La investigación se orienta a conocer la característica del *Hardware* que sea óptimo para la implementación del modelo requerido.
- *Software*. El *Software* debe permitir implementar los requisitos funcionales y permitir requisitos funcionales a nivel del estado del arte para este tipo de soluciones.

En la industria de *Machine Learning* y *Deep Learning* existen las siguientes alternativas de *Hardware*:



Figura 5. Alternativas de HW para soluciones de ML y DL (Jawandhiya, 2018).

Las *CPUs* son las unidades de procesamiento de nuestros computadores y teléfonos móviles. En estos dispositivos podemos usar múltiples lenguajes de programación. Las *GPUs*, unidades de procesamiento gráfico, son diseñadas para procesar tareas en paralelo, en general, las *GPUs* se pueden desempeñar bien ejecutando algoritmos de *Deep Learning* porque requieren un alto número de cálculos como sumas y multiplicaciones (Jangamreddy, 2019). Los *FPGAs* (matrices de puertas programables por campo) son un tipo de *hardware* que puede ser programado y reconfigurado usando *HDL* (Lenguaje descriptivo de *Hardware*). Se ha observado que la relación de desempeño sobre energía consumida es más alta que para los *GPUs*, y la flexibilidad del *hardware* puede tener ventajas a futuro con los modelos de *Machine Learning* (*ML*) y *Deep Learning* (*DL*) que también son cambiantes (Jawandhiya, 2018). Por esta razón, varias empresas han usado estos dispositivos como soporte para su estrategia de servicios de *ML* y *DL*, como veremos más adelante. También existen los *ASICs* (circuitos integrados específicos de aplicación). Es la opción menos flexible, pero a la vez tiene el desempeño más alto de las alternativas (Jawandhiya, 2018). *Google* usa este tipo de implementaciones con su *TPU* (*Tensor Processor Unit*) (Stone, 2019).

En pruebas realizadas con modelos de *DNN* (Redes neuronales artificiales profundas) se identificó que un *TPU* usa menos de la mitad de la potencia de un sistema de *GPU* y tiene mejor latencia, incluso alcanzando tiempos de respuesta 15 veces más rápidos en el procesamiento que un sistema *GPU*, lo cual hace que *TPU* se convierta en un arquetipo

para arquitecturas específicas de dominio (Jouppi, et al., 2017). El tiempo de latencia es muy importante para la experiencia del usuario en la aplicación a implementar.

Teniendo en cuenta lo anterior, para usar los modelos de nivel de estado del arte es conveniente usar servicios en la nube. En los siguientes capítulos, se revisan las alternativas de los proveedores de servicios de nube más destacados para implementar la solución que tiene por objetivo este trabajo.

El objetivo del presente trabajo es encontrar solución a las **necesidades corporativas** expuestas para organizaciones que hagan uso de información masiva de clientes, o usuarios, o ciudadanos. Otro objetivo es usar los modelos de análisis de sentimiento de mejor resultado comprobados por la **academia** para suplir estas necesidades, usando las mejores **arquitecturas, técnicas y herramientas tecnológicas** para asegurar el cumplimiento de requisitos de análisis de información priorizando los siguientes aspectos:

- Buen desempeño de herramientas.
- Facilidad de uso.
- Resultados comprobables y satisfactorios.

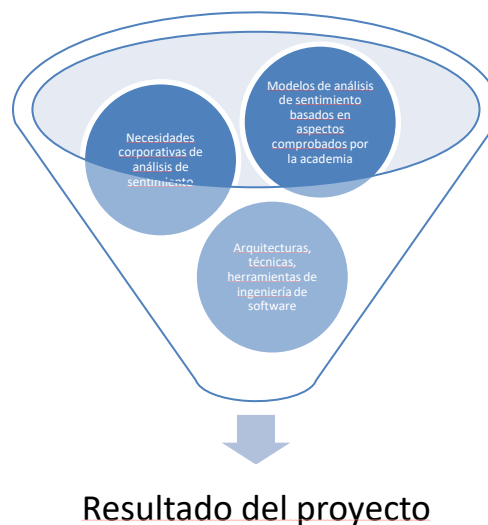


Figura 6. *Elementos de entrada para el proyecto propuesto* (elaboración propia).

El objetivo del presente trabajo no es entrar en detalle en la operación de los modelos basados en redes neuronales debido a que es un trabajo de ingeniería de *software*. Para el presente trabajo consideramos los modelos a usar como “cajas negras” para articular las herramientas corporativas que van a hacer uso del análisis para el apoyo en la orientación de decisiones.

2.2 Conjunto de soluciones complementarias

A continuación se analiza la solución propuesta por los principales proveedores de servicios de computación en la nube para dar soporte al entrenamiento y uso del modelo *BERT*.

Google Cloud Platform





A nivel de *Hardware* se propone el *TPU* versión 3:



Figura 7. Hardware propuesto por Google Cloud Platform para modelos de ML y DL – Cloud TPU v3 (Sato, 2018).

Esta solución está optimizada para procesar cargas de los siguientes modelos:

Wide range of reference models for Cloud TPUs

 <p>Image recognition, segmentation, & more</p> <p>Image Recognition: AmoebaNet-D ResNet-50/101/152/200 Inception v2/v3/v4</p> <p>Object Detection: RetinaNet, Mask R-CNN</p> <p>Image Segmentation: Mask R-CNN, DeepLab, and RetinaNet</p> <p>Low-Resource Models: MobileNet SqueezeNet</p>	 <p>Machine translation and language modeling</p> <p>Machine translation Language modeling Sentiment analysis (all Transformer-based)</p> <p>Question-answering (QANet)</p> <p>BERT: State-of-the-art results across 13 natural language tasks</p>	 <p>Speech recognition</p> <p>ASR Transformer (LibriSpeech)</p>	 <p>Image generation</p> <p>Image Transformer DCGAN</p>
--	---	---	---

g.co/cloudtpu

Figura 8. Modelos de ML y DL para Cloud TPUs (Sato, 2018).

Entre los modelos listados se encuentra *BERT* y un modelo de “*Speech recognition*” para pasar voz a texto, para hacer el análisis de información de audios.

Para la configuración de los servicios en la nube, sintonía fina de modelo *BERT*, despliegue en producción y uso de modelo, en *Google Cloud Platform* se sugieren los siguientes pasos (Google, 2020c):

1. Crear una variable para el código del proyecto.
2. Configurar la línea de comando para usar el proyecto donde se desee crear el *Cloud TPU*.
3. Crear un *Cloud Storage Bucket* para guardar los datos que se usan para entrenar el modelo y los resultados del proceso de sintonía fina.
4. Establecer permisos del *Cloud TPU* al *Cloud Storage Bucket*.
5. Iniciar un *Compute Engine VM* (Máquina virtual) y el *Cloud TPU*.
6. Definir variables de ambiente.
7. Clonar el repositorio de *BERT* en *Github*.
8. Descargar datos de sintonía fina y evaluación en formato *TfRecord*.
9. Entrenar el modelo.
10. Verificación de resultados.
11. Ajuste de hiperparámetros.
12. Despliegue del modelo.
13. Uso en datos de producción.

De manera alternativa *Google* ofrece la solución *Colab* para facilitar la ejecución de tutoriales con la edición y ejecución de código *Python* y *Tensorflow* (Google, 2019). Se siguen los pasos mencionados anteriormente para el uso de *BERT* y *BETO*:

- *BERT* - Tareas de clasificación de oraciones y pares de oraciones (Google, 2018).
- *BETO* - Tarea de completar oraciones (Cañete, et al., 2020b)

Microsoft Azure

A nivel de *hardware* se propone el uso de *FPGAs* o *GPUs*. Para *FPGAs* se requiere convertir los modelos de Tensorflow a formato *ONNX* (*Open Neural Network Exchange*) (Microsoft, 2020d). Para *FPGA* están disponibles modelos entre los cuales no se incluye *BERT*. Sin embargo, Microsoft hizo una contribución en *ONNX* para mejorar el desempeño del modelo *BERT* en una escala de 17x usando *hardware* de tipo *CPU* y *GPU* (Alford, 2020).

Para el uso de *BERT* se propone la siguiente secuencia de actividades con el uso de *GPUs* (Zhu, 2020):

1. Inicializar espacio de trabajo de *Azure Machine Learning*.
2. Descargar modelo y datos de sintonía fina en Azure Blob Container.
 - a. Crear directorio de proyecto.
 - b. Clonar el repositorio de *Github* de *BERT*.
 - c. Definir el almacén de datos de Azure ML para guardar el archivo del modelo convertido.
 - d. Crear un experimento de *Azure ML* para almacenar todas las ejecuciones.
 - e. Crear un estimador de *Tensorflow*.
 - f. Iniciar y monitorear la ejecución.
3. Sintonía fina con *BERT* y datos de prueba (A nivel de *hardware* se usa *GPU*)
 - a. Crear un estimador de *Tensorflow* para sintonía fina.
 - b. Iniciar y monitorear ejecución.
4. Sintonía fina de *BERT* con hiperparámetros
 - a. Inicio de barrido de hiperparámetros.
 - b. Monitorear las ejecuciones de sintonía de hiperparámetros.
 - c. Encontrar y registrar el mejor modelo para el uso con datos productivos.

AWS – Amazon web services

Amazon propone el uso de su producto *SageMaker* para hacer sintonía fina del modelo *BERT* pre-entrenado en infraestructura de *GPUs* usando pasos similares a los descritos

para *Google Cloud Platform* y Microsoft Azure (Li & Ping, 2020). AWS tiene soporte de *hardware FPGAs* aunque no para uso de modelo *BERT* (Amazon, 2020).

Otros proveedores de *cloud computing*

En *IBM Cloud* y *Alibaba* se encuentran ejemplos de *BERT* preconstruidos para su uso, no se entra en el detalle de la implementación del proceso de sintonía fina. Incluso Alibaba propone *Perseus BERT*, una solución para optimizar *BERT* pasada en *GPUs* (Lixiang, Zhigang, Zhiming, & Liang, 2019).

Teniendo en cuenta lo anterior, debido al origen de *BERT*, y su optimización en *Hardware*, las tareas de planificación se plantearán con base en la propuesta de *Google Cloud Platform*.

2.3 Conjunto de soluciones similares o parciales

En general, las propuestas de soluciones similares no incluyen los siguientes aspectos:

- Explicación de modelo usado, o alguna evaluación de *benchmark* tal como *SemEval*.
- Solución de análisis de sentimiento **basada en aspectos**. Algunas son soluciones que evalúan el sentimiento de la frase completa.

A continuación se listan por separado soluciones de análisis de sentimiento de audios de *call center* y *APIs* de análisis de sentimiento basado en aspectos. Se investigaron soluciones que analizan lenguaje natural en idioma español.

Soluciones de análisis de sentimiento de audios de *call center* en idioma español

- *Vonage*: Se ofrece análisis de sentimiento en línea por medio de un “*bot*” para generar alertas. Con los resultados el *bot* se entrena para mejorar su análisis (Vonage, 2020).

El proceso propuesto es el siguiente:



Figura 9. Solución de análisis de sentimiento propuesta por Vonage (Vonage, 2020).

El proyecto propuesto en el presente trabajo presenta la alternativa de hacer más transparente la operación y mejora del modelo que analiza la información. No se prioriza la respuesta en línea sino el análisis general del proceso de servicio al cliente para implementar mejoras en el proceso integral.

- *Advanced analytics*: Se menciona proceso y análisis de textos y análisis de audios pero no se da mayor detalle del proceso sugerido (Apex, 2020). Esta empresa es *partner* de *Google Cloud Platform*.

APIs de análisis de sentimiento en español

Las siguientes APIs tienen soporte para análisis de modelo basado en aspectos en idioma español, pero no explican que modelo usan para evaluar la efectividad del modelo sobre un *benchmark* como *SemEval*:

- *Cloudmersive* (Cloudmersive, 2020). Tiene una ventaja que puede ser usado por un conector de *PowerApps*. En la prueba gratuita se puede ver el response y request del servicio de análisis de sentimiento:

```

SentimentAnalysisRequest {
  description: Input to a sentiment analysis operation
  TextToAnalyze string
  Text to analyze
}

SentimentAnalysisResponse {
  description: Output of a sentiment analysis operation
  Successful boolean
  True if the sentiment analysis operation was successful, false otherwise

  SentimentClassificationResult string
  Classification of input text into a sentiment classification; possible values are "Positive", "Negative" or "Neutral"

  SentimentScoreResult number($double)
  Sentiment classification score between -1.0 and +1.0 where scores less than 0 are negative sentiment, scores greater than 0 are positive sentiment and scores close to 0 are neutral. The greater the value deviates from 0.0 the stronger the sentiment, with +1.0 and -1.0 being maximum positive and negative sentiment, respectively.

  SentenceCount integer($int32)
  Number of sentences in input text
}

```

Figura 10. Request y response de API de sentimiento de Cloudmersive (Cloudmersive, 2020).

La API toma una frase completa y entrega como salida clasificación y puntaje de sentimiento. La API no maneja la evaluación por aspectos y no da la opción de entrenar el modelo.

- *Meaning Cloud* (MeaningCloud, 2020)

Meaning Cloud propone el uso de varias herramientas para el análisis de sentimiento:

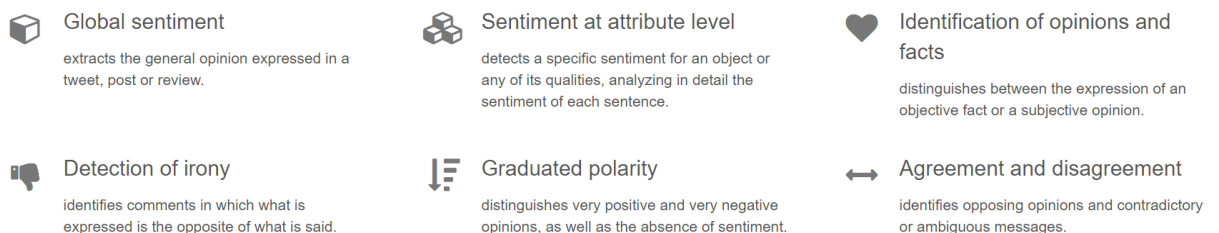


Figura 11. Herramientas de análisis de sentimiento de Meaning Cloud (MeaningCloud, 2020)

La API de Análisis de Sentimiento cuenta con la posibilidad de analizar atributos como se ve en la siguiente figura extraída de un webinar de Meaning Cloud.

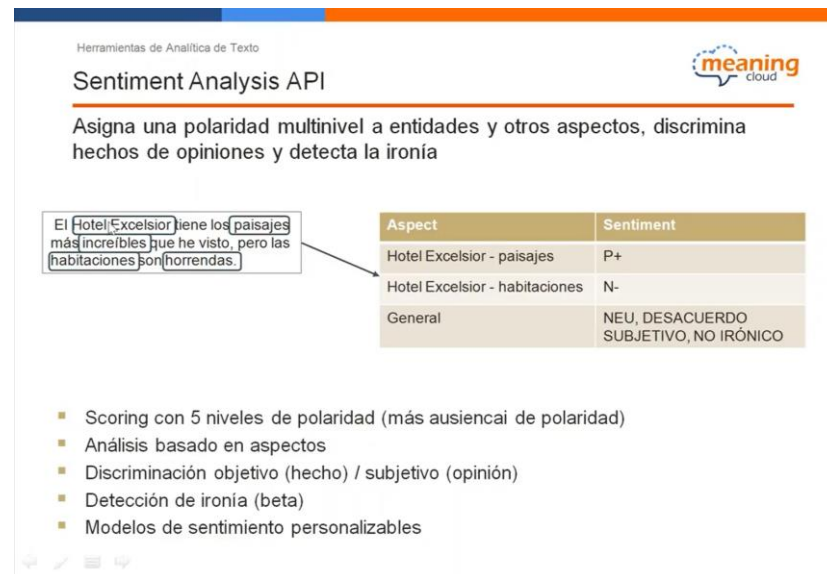
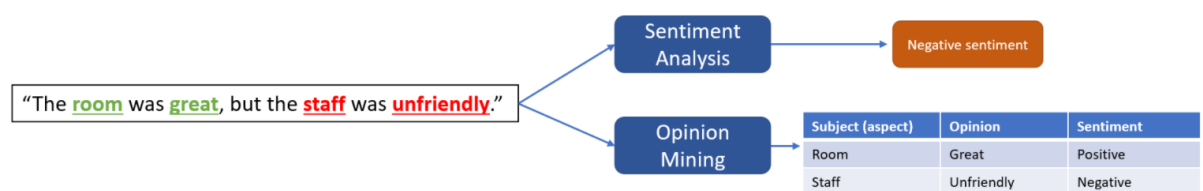


Figura 12. Identificación de análisis de sentimiento basado en aspectos de la API de Meaning Cloud (Matarranz, 2017)

Por la fecha del webinar puede que se usen redes neuronales de memoria corta, no se tiene detalle del modelo usado y el proceso de entrenamiento para definir las entidades que definen los aspectos.

- Microsoft sentiment analysis (Steen, et al., 2020)

En esta API también se tiene la opción de identificar el análisis de sentimiento por aspectos como se ve en la siguiente imagen.



A diagram of the Opinion Mining example

Figura 13. Identificación de análisis de sentimiento basado en aspectos de la API de análisis de sentimiento de Microsoft (Steen, et al., 2020)

Tal como la anterior API, no se detalla el modelo usado y el proceso de entrenamiento, funciona como caja negra.

- Repustate (Repustate, 2020)

En ésta API es posible analizar sentimiento basado en aspectos, hay que enviarle los aspectos a analizar en el campo “*topics*” como se ve en la siguiente figura.

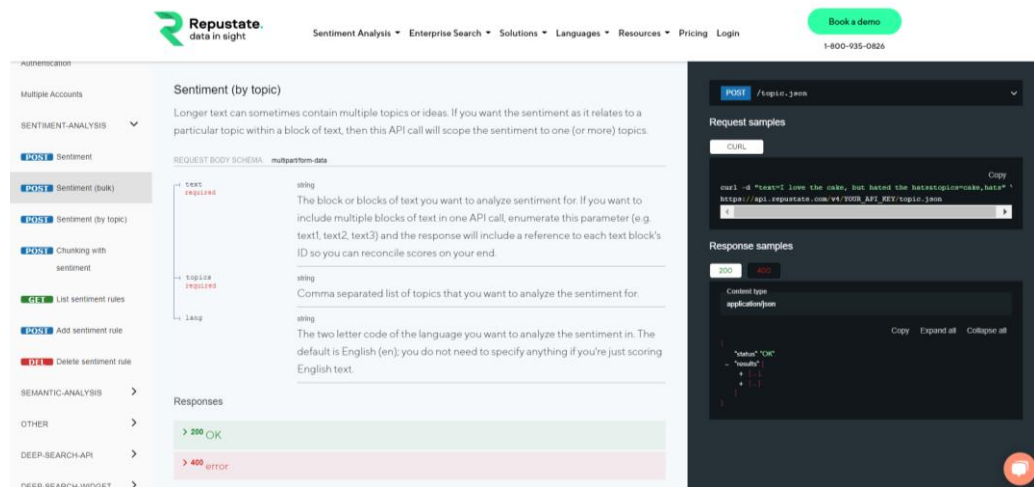


Figura 14. Identificación de análisis de sentimiento basado en aspectos de la API de Repustate (Repustate, 2020)

En el presente proyecto el objetivo es que en el proceso de sintonía fina del modelo se definan los aspectos con base en los datos de entrada para determinarlos a nivel general, sin que haya necesidad de especificarlos en cada *request*.

2.4. Conclusiones

El estado del arte inicia evaluando alternativas de herramientas de procesamiento de lenguaje natural para el objetivo del trabajo, se describen resultados de esfuerzos anteriores de modelos basados en redes neuronales artificiales, se analizan soluciones a nivel de *software* y *hardware*, y análisis de soluciones parciales. Hasta ahora no se ha encontrado una solución integral que tenga las siguientes características:

- Basada en un modelo de nivel de estado del arte.
- Uso de análisis de sentimiento basado en aspectos.
- Multilenguaje o en idioma español.

Las APIs de Meaning Cloud y Microsoft podrían ser interesantes para usar, aunque los parámetros de entrenamiento del modelo no son manejables. Por lo cual se va a plantear la planificación de sintonía fina de modelo *BERT*, con consumo de *API* de paso de voz a texto e integración de resultados con indicadores convencionales de análisis.

3. Objetivos concretos y metodología de trabajo

3.1. Objetivo general

El objetivo del proyecto es automatizar el análisis de sentimiento basado en aspectos usando *APIs* y modelos pre-existentes de procesamiento de lenguaje natural para clasificar información e implementar flujos de trabajo para entornos corporativos para reducir la carga operativa de tareas y aumentar el porcentaje de dedicación de los usuarios corporativos a tareas de análisis.

3.2. Objetivos específicos

- Validar en fuentes académicas el tipo de modelo a usar para el análisis de sentimiento basado en aspectos como una tarea de procesamiento de lenguaje natural. El modelo debe haber sido probado en tareas similares, y permitir la adaptación a un contexto particular.
- Revisar en los proveedores de servicios de computación en la nube las mejores opciones a nivel de *hardware* y *software* para la implementación del modelo.
- Planificar el desarrollo de un flujo de trabajo de análisis de información que use los resultados del modelo para la toma de decisiones estratégicas sobre el proceso de atención a comunicaciones de usuarios.
- Establecer un proceso organizacional complementario a la funcionalidad del flujo de trabajo para facilitar el cambio para el uso de técnicas novedosas.
- Planificar la toma de requisitos de clientes para la adaptación de modelo a necesidades particulares a nivel de aspectos, terminología, y flujo de trabajo.

3.3. Metodología del trabajo

Se hace uso de técnicas ágiles para el desarrollo de las fases del proyecto. Aunque se espera mantener un fácil entendimiento de la solución y su arquitectura interna, su adaptación a un contexto particular puede necesitar del levantamiento de requisitos personalizados que resulte en zonas de incertidumbre que se van aterrizando con las entregas continuas de los *sprints*. En el plan se proponen las actividades con base en *sprints* de 2 semanas. En el desarrollo de los *sprints* se manejan las reuniones propuestas por *Scrum*, la reunión diaria, planificación, revisión y retrospectiva. Las estimaciones se

realizan con el equipo de desarrollo con la aplicación *Scrum Poker Cards*. La unidad de estimaciones se lleva en horas. El control del tiempo tomado en las tareas y las asignaciones se manejan en JIRA. De JIRA se consulta el *Burndown Chart* para llevar el control diario del avance y cumplimiento.

Para el desarrollo del *Workflow* corporativo se propone el uso de *Microsoft PowerApps* ya que facilita la elaboración de *mockups* teniendo entregas tempranas e incluso adaptaciones personalizadas manejadas por un desarrollador experto (Microsoft, 2020c).

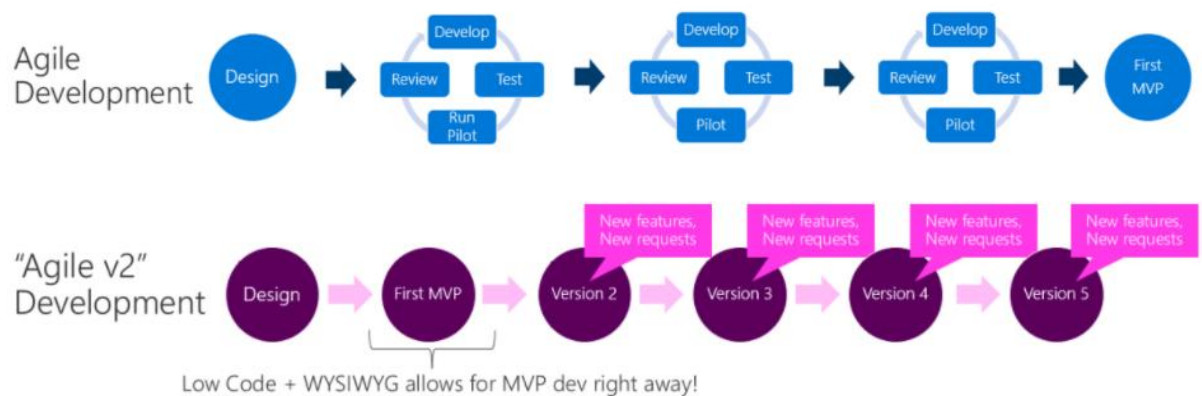


Figura 15. Técnica *Agile v2* propuesta por framework *Power Apps* (Microsoft, 2020c).

Para la sintonía fina del modelo, en el siguiente capítulo se dará el detalle de las fases para consolidar y codificar los datos de prueba, los cuales son claves para la adaptación a un contexto particular de la solución.

4. Desarrollo específico de la contribución

4.1. Estructura de desglose de trabajo

Se propone dividir la estructura de desglose de trabajo por los diferentes módulos de la solución propuesta. Los módulos son los siguientes:

- Proceso de sintonía fina para personalización de modelo para análisis de las variables requeridas para el negocio: En este proceso se recaba la información de negocio y se obtiene un modelo con sintonía fina teniendo en cuenta los datos de prueba personalizados para identificar el análisis de sentimiento de los aspectos claves para el negocio.
- Proceso de paso de voz a texto e identificación de roles: En este proceso se pasa la información de audios de atenciones a clientes a textos, identificando los roles de cada participante en la conversación.
- Proceso de sintonía fina frecuente de modelo: Este proceso permite continuar entrenando el modelo teniendo en cuenta los resultados del primer esfuerzo de implementación.
- Proceso de consolidación de análisis: Es el proceso de análisis de información productiva para obtener análisis lo más cercano posible al tiempo real.
- Proceso de validación de información: Es el proceso manejado por los analistas de negocio para identificar patrones para tomar acciones de gestión en el modelo de servicio a nivel del personal de servicio al cliente y las demás áreas que componen la cadena de valor de la organización.
- Proceso de gestión de información: Una vez identificada la información del análisis de sentimiento basado en aspectos se procede a tomar acciones de gestión en el personal de servicio y gestión en las áreas que componen la cadena de valor de la organización para implementar un modelo de mejora continua del proceso integral.
- Proceso de análisis ejecutivo: En este proceso se analizan indicadores de cumplimiento de los anteriores procesos. El objetivo es orientar la estrategia en función del análisis del sentimiento controlando las variables de cada uno de los procesos operativos que mantienen operando el modelo.

La herramienta que se planifica como objetivo de este trabajo debe dar soporte a todos los procesos corporativos que hacen parte del análisis:

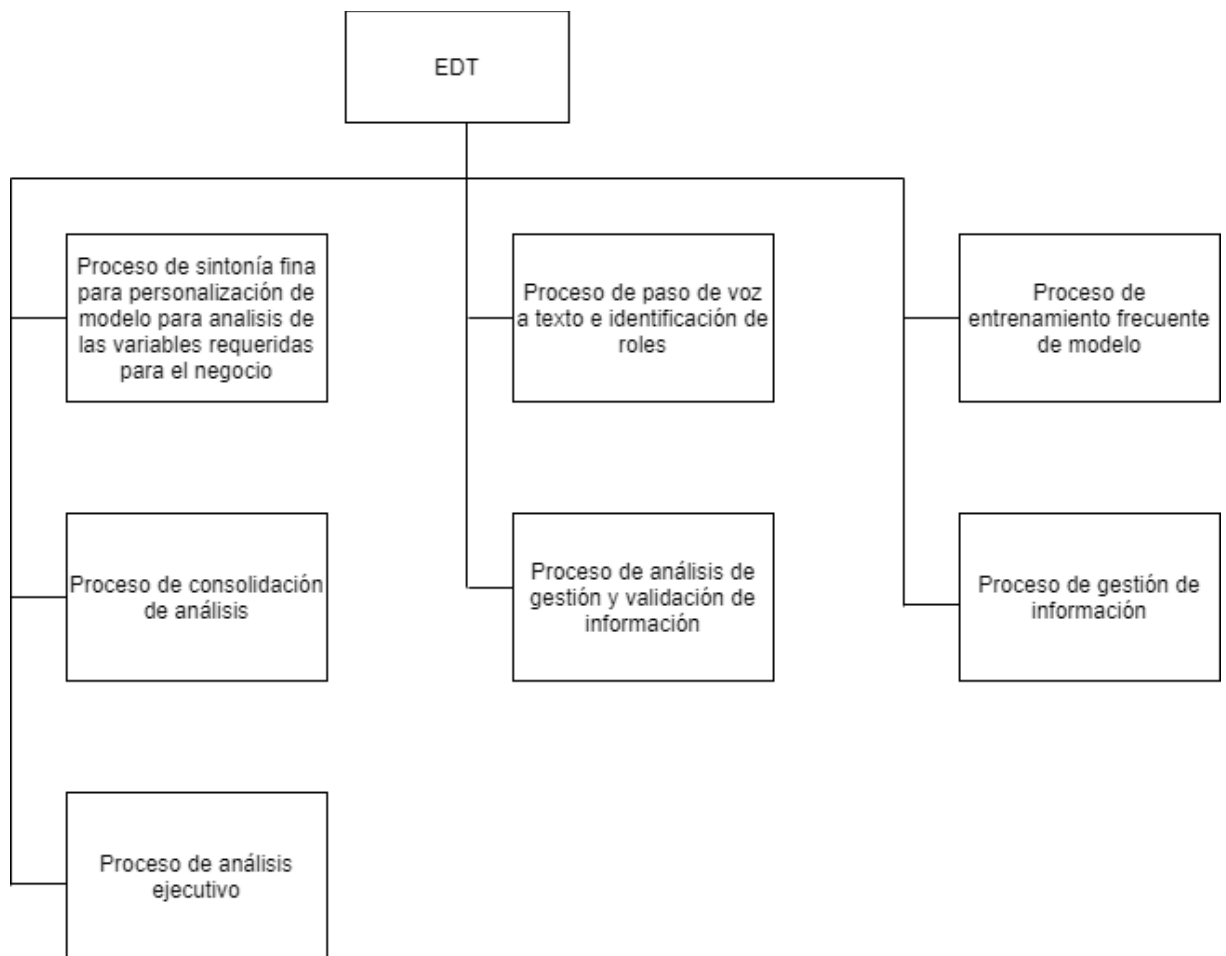


Figura 16. Estructura de desglose de trabajo de proyecto (elaboración propia).

4.2. Estudio del dominio

Uno de los objetivos de esta fase es definir los aspectos importantes de análisis para un contexto particular. A continuación un ejemplo de aspectos para una entidad bancaria:



Figura 17. *Ejemplo de aspectos de análisis de sentimiento para la industria bancaria* (Lappemana, Clark, Evans, Sierra-Rubiac, & Gordon, 2020)

En adelante se van a suponer indicadores estándar de la industria de servicio al cliente para la evaluación de sentimiento que por lo general se miden por encuestas:

- Satisfacción.
- Atención.
- Resolución.
- Calidad.
- Experticia.

En los análisis se van a complementar estos indicadores con los convencionales de la industria:

- Abandono.
- Tiempo de espera.
- Tiempo de servicio.

4.3. Análisis y modelado de requisitos de *software*

Para iniciar el análisis de los requisitos, se detalla un proceso en el que se detallan los roles que van a hacer uso de la herramienta para alcanzar los objetivos corporativos.

4.3.1. Proceso sugerido

Se sugiere la implementación de los siguientes procesos:

- Proceso de consolidación de análisis.
- Proceso de validación de información.
- Proceso de gestión de información analizada.
- Proceso de sintonía fina continua de modelo.
- Proceso de análisis ejecutivo.

4.3.2. Proceso de consolidación de análisis

El proceso de consolidación controla los pasos de procesamiento que debe tener una grabación de atención al usuario para:

- pasar a texto,
- a formato *TF-Record*
- y ejecución de análisis de sentimiento basado en aspectos.

El usuario de este proceso, o “Analista de consolidación de análisis” debe tener conocimiento en todos los procesos de información para monitorear y dar soporte en línea al proceso completo. También debe tener acceso a la interfaz que permite cambiar los recursos disponibles para el procesamiento de información.

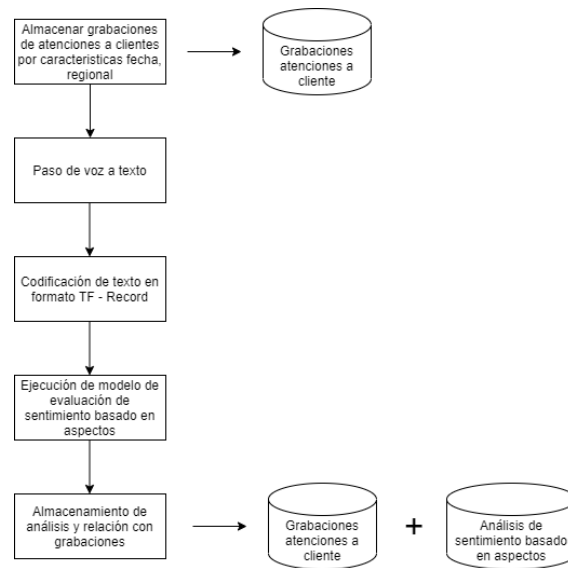


Figura 18. *Proceso de consolidación de análisis basado en aspectos* (Elaboración propia)

4.3.3. Proceso de validación de información

El “Analista de validación de información” puede ser un analista asignado para revisar el análisis de sentimiento de un servicio y una regional específica, o un analista ejecutivo que puede tener una perspectiva global y ver todas las regionales, servicios, e información sensible de clientes. Este usuario cuenta con acceso a las perspectivas analíticas y a la reproducción de grabaciones de atenciones específicas.

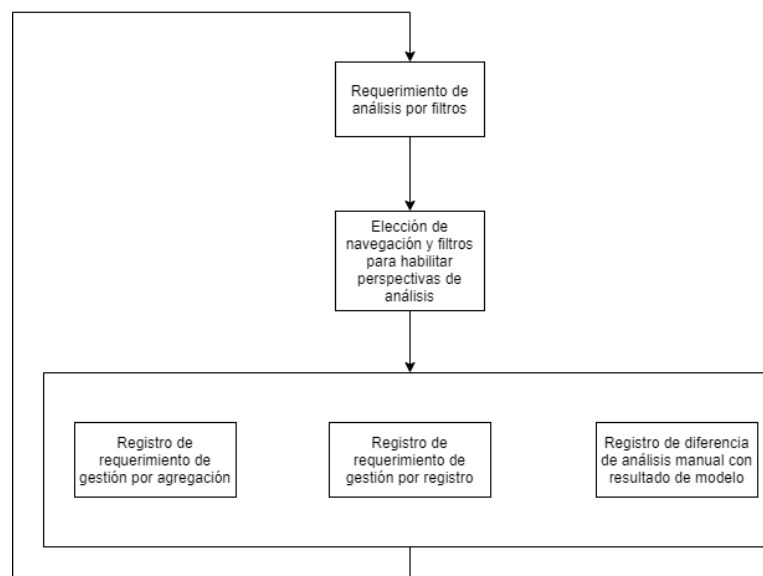


Figura 19. *Proceso de analista de validación y gestión de información* (Elaboración propia)

4.3.4. Proceso de gestión de información analizada

El proceso es manejado por el rol de “Analista de gestión”. Este rol con base en las variables generales o detalladas identificadas en el proceso de validación es responsable de realizar las siguientes acciones:

- Gestión de procesos de servicio.
- Capacitación a personal de servicio.
- Acciones de personal.
- Gestión con áreas que componen la cadena de valor de la organización.

4.3.5. Proceso de sintonía fina de modelo

El proceso inicial de sintonía fina requiere del levantamiento de requisitos de negocio, y las actividades propias de la sintonía fina del modelo:

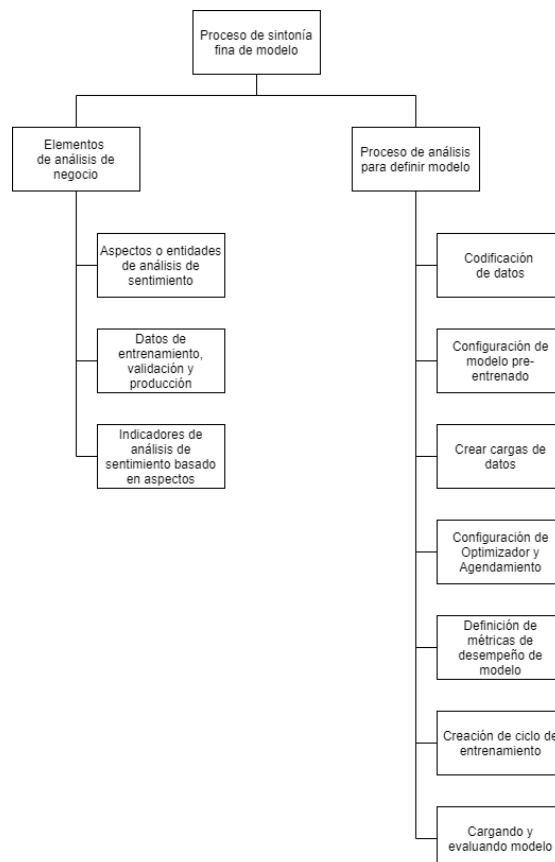


Figura 20. Proceso de sintonía fina de modelo inicial (Google, 2020c)

Una vez el modelo está en producción, se requiere implementar un proceso de sintonía fina frecuente alimentado por la validación manual de analistas. El proceso cuenta con las siguientes fases:

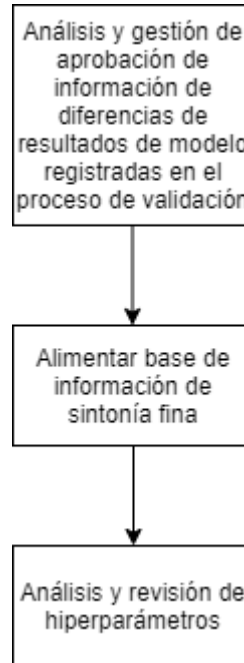


Figura 21. *Proceso de sintonía fina frecuente de modelo (Google, 2020c)*

4.3.6. Proceso de análisis ejecutivo

Un “Analista ejecutivo” puede realizar las siguientes acciones:

- Análisis de sentimiento basado en aspectos global o segmentado por servicio, regional, o segmento de cliente.
- Análisis proceso de consolidación de información.
- Análisis proceso de validación.
- Análisis proceso de gestión.
- Análisis de proceso de entrenamiento de modelo.

4.4. Planificación

4.4.1. *Visual Story Mapping*

Con el objetivo de mejorar la visualización del *Visual Story Mapping* se divide por los diferentes módulos o fases del proyecto con la siguiente codificación de colores:

- Gris para los módulos o fases.
- Azul para las historias de usuario.
- Verde para las tareas:

La primera parte contiene historias y tareas del levantamiento de requisitos. En esta fase es importante la planificación de grupos focales con diversos actores para entender los aspectos de análisis como insumo a los datos que van a consolidarse para hacer la sintonía fina del modelo.

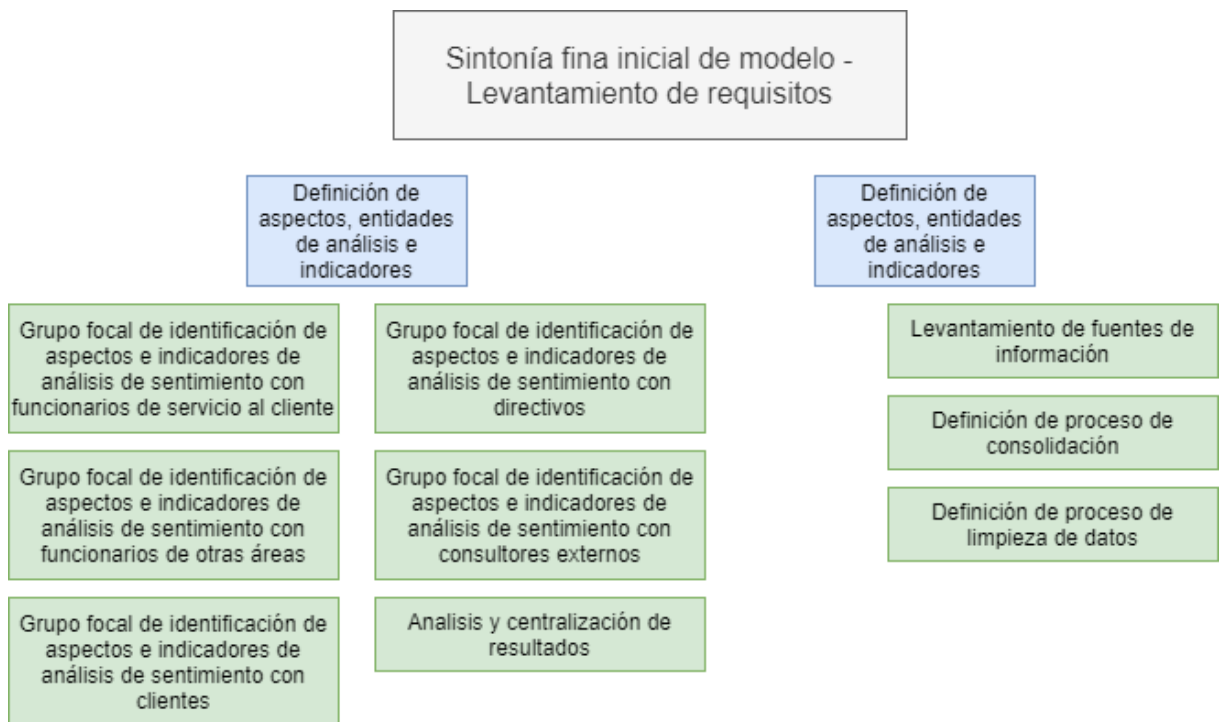


Figura 22. *Visual Story Mapping* – Fase de Sintonía fina inicial de modelo – Levantamiento de requisitos (Elaboración propia)

En esta parte están historias y tareas de análisis y diseño. El análisis realizado sobre los datos de sintonía fina y procesamiento son determinantes para los resultados del modelo.

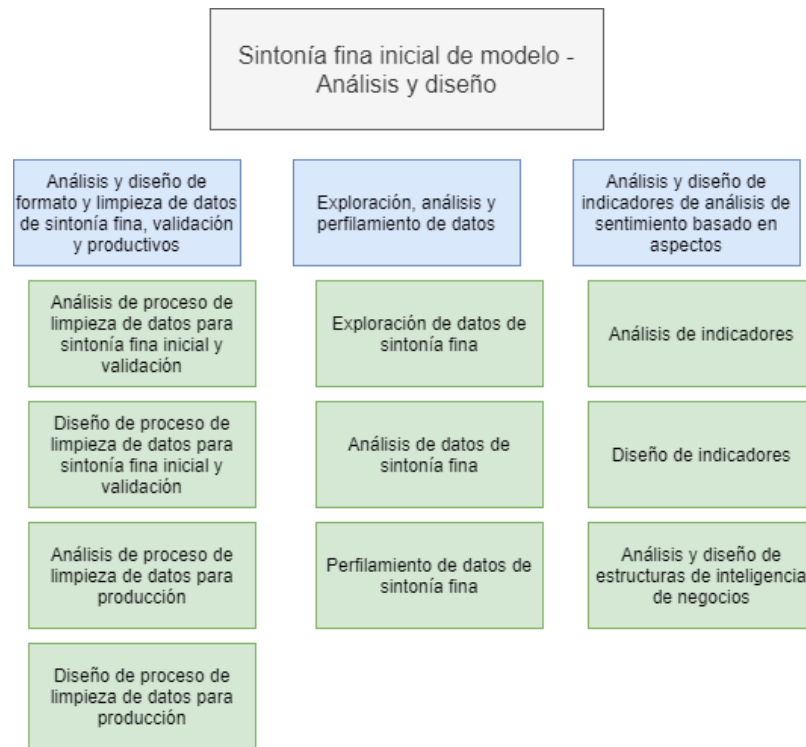


Figura 23. *Visual Story Mapping – Fase de Sintonía fina inicial de modelo – Análisis y diseño*
(Elaboración propia)

En la siguiente parte se inicia con tareas de desarrollo y pruebas y en paralelo se ajusta el modelo de análisis de sentimiento basado en aspectos para iniciar a usarlo.

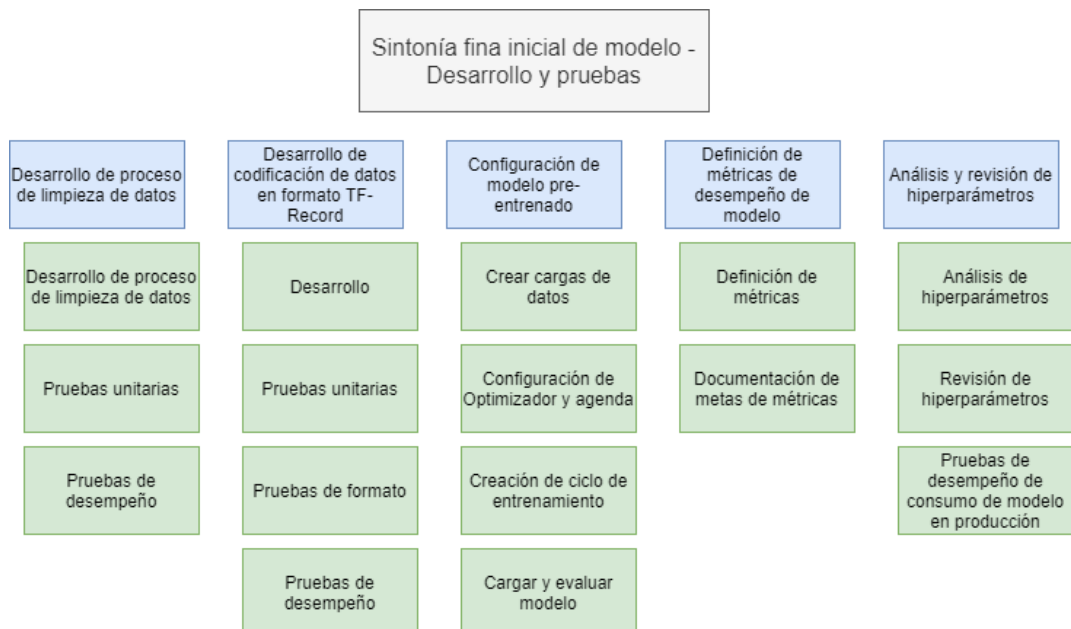


Figura 24. *Visual Story Mapping – Fase de Sintonía fina inicial de modelo – Desarrollo y pruebas*
(Elaboración propia)

En esta parte se tienen las tareas de pruebas de consumo de APIs de paso de voz a texto y del servicio productivo de análisis de sentimiento y en paralelo el desarrollo de la orquestación del proceso de información.

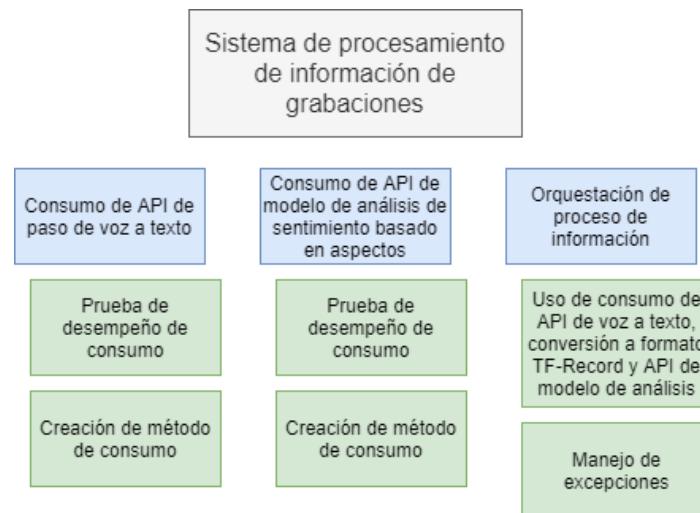


Figura 25. *Visual Story Mapping – Modulo de sistema de procesamiento de información de grabaciones* (Elaboración propia)

En esta parte se tiene el desarrollo de las estructuras de análisis que se van a usar en la consolidación de analíticas para validación.

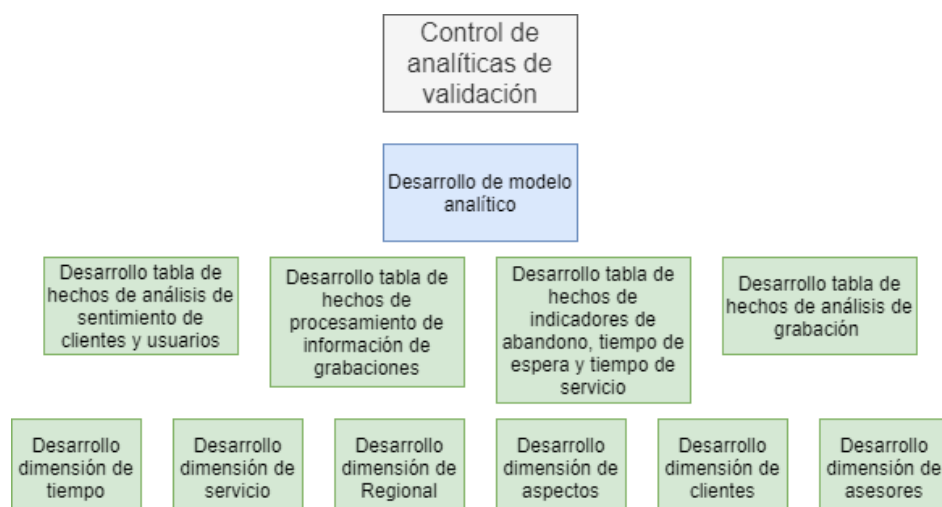


Figura 26. *Visual Story Mapping – Modulo de control de analíticas de validación* (Elaboración propia)

En la siguiente parte se tiene el desarrollo de los módulos componentes complementarios, Workflow de gestión, control de roles. También están las pruebas integrales y las encuestas de validación de usabilidad.

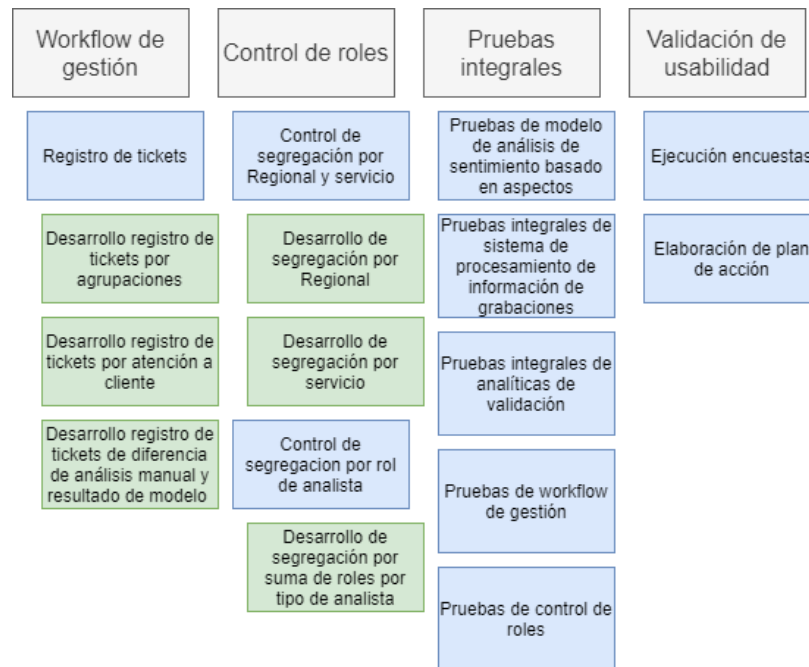


Figura 27. *Visual Story Mapping – Módulos de Workflow de gestión, control de roles y Fases de Pruebas integrales y Validación de usabilidad* (Elaboración propia)

En esta sección se determinan los módulos y fases necesarias para la ejecución del proyecto con sus historias y tareas, sin determinar secuencia ni tiempos. La secuencia de actividades, tiempos y esfuerzos se determinarán en el calendario de ejecución.

4.4.2. Distribución de esfuerzos

Para el desarrollo de las actividades detalladas en el *Visual Story Mapping* se requiere contar con profesionales con los siguientes roles:

- *Product Owner*
- Ingeniero de datos
- Diseñador de soluciones de *Business Intelligence*
- Desarrollador de *Backend* e integración de servicios

- Desarrollador de *Business intelligence*
- Desarrollador de *Frontend*
- Analista de pruebas
- *Scrum master*

En el calendario del siguiente numeral se relacionarán las actividades identificadas en el *Visual Story Mapping* con los responsables y tiempos.

4.4.3. Calendario de ejecución

La planificación se hace en base a *sprints* de 2 semanas, y las estimaciones en horas:

Tabla 3. Calendario de ejecución

Módulo / Fase	Historia de usuario	Tarea	Responsable	Horas	Sprint
		Sintonía fina inicial de modelo - Levantamiento de requisitos			
		Definición de aspectos, entidades de análisis e indicadores			
		Grupo focal de identificación de aspectos e indicadores de análisis de sentimiento con funcionarios de servicio al cliente	Product Owner	3	1
		Grupo focal de identificación de aspectos e indicadores de análisis de sentimiento con funcionarios de otras áreas	Product Owner	3	1
		Grupo focal de identificación de aspectos e indicadores de análisis de sentimiento con clientes	Product Owner	3	1
		Grupo focal de identificación de aspectos e indicadores de análisis de sentimiento con directivos	Product Owner	3	1
		Grupo focal de identificación de aspectos e indicadores de análisis de sentimiento con consultores externos	Product Owner	3	1
		Análisis y centralización de resultados	Product Owner	2	1
		Definición de datos de sintonía fina, validación y producción			
		Levantamiento de fuentes de información	Product Owner	6	1
		Definición de proceso de consolidación	Product Owner / Desarrollador de Backend	6	1
		Definición de proceso de limpieza de datos	Product Owner / Desarrollador de Backend	6	1
		Sintonía fina inicial de modelo - Análisis y diseño			
		Análisis y diseño de formato y limpieza de datos de sintonía fina, validación y productivos			
		Análisis de proceso de limpieza de datos para sintonía fina inicial y validación	Ingeniero de datos	16	2
		Diseño de proceso de limpieza de datos para sintonía fina inicial y validación	Ingeniero de datos	16	2
		Análisis de proceso de limpieza de datos para producción	Ingeniero de datos	16	2
		Diseño de proceso de limpieza de datos para producción	Ingeniero de datos	16	2

Módulo / Fase	Historia de usuario	Tarea	Responsable	Horas	Sprint
		Exploración, análisis y perfilamiento de datos			
		Exploración de datos de sintonía fina	Ingeniero de datos	16	2
		Análisis de datos de sintonía fina	Ingeniero de datos	16	2
		Perfilamiento de datos de sintonía fina	Ingeniero de datos	16	2
		Análisis y diseño de indicadores de análisis de sentimiento basado en aspectos			
		Análisis de indicadores	Product Owner / Diseñador solución <i>Business Intelligence</i>	8	2
		Diseño de indicadores	Product Owner / Diseñador solución <i>Business Intelligence</i>	16	2
		Análisis y diseño de estructuras de Business Intelligence	Diseñador solución <i>Business Intelligence</i>	16	2
		Sintonía fina inicial de modelo - Desarrollo y pruebas			
		Desarrollo de proceso de limpieza de datos			
		Desarrollo de proceso de limpieza de datos	Desarrollador de <i>Backend</i>	32	3
		Pruebas unitarias	Desarrollador de <i>Backend</i> / Analista de pruebas	16	3
		Pruebas de desempeño	Desarrollador de <i>Backend</i> / Analista de pruebas	16	3
		Desarrollo de codificación de datos en formato <i>TF-Record</i>			
		Desarrollo	Desarrollador de <i>Backend</i>	16	3
		Pruebas unitarias	Desarrollador de <i>Backend</i> / Analista de pruebas	16	3
		Pruebas de formato	Desarrollador de <i>Backend</i> / Analista de pruebas	16	3
		Pruebas de desempeño	Desarrollador de <i>Backend</i> / Analista de pruebas	16	3
		Configuración de modelo pre-entrenado			
		Crear cargas de datos	Ingeniero de datos	8	3
		Configuración de Optimizador y agenda	Ingeniero de datos	8	3
		Creación de ciclo de entrenamiento	Ingeniero de datos	8	3
		Cargar y evaluar modelo	Ingeniero de datos	16	3
		Definición de métricas de desempeño de modelo			
		Definición de métricas	Ingeniero de datos	8	3
		Documentación de metas de métricas	Ingeniero de datos	16	3
		Análisis y revisión de hiperparámetros			
		Análisis de hiperparámetros	Ingeniero de datos	16	3
		Revisión de hiperparámetros	Ingeniero de datos	16	3
		Pruebas de desempeño de consumo de modelo en producción	Ingeniero de datos	16	3
		Sistema de procesamiento de información de grabaciones			
		Consumo de <i>API</i> de paso de voz a texto			
		Prueba de desempeño de consumo	Desarrollador de <i>Backend</i> e integración	8	4
		Creación de método de consumo	Desarrollador de <i>Backend</i> e integración	8	4

Módulo / Fase	Historia de usuario	Tarea	Responsable	Horas	Sprint
		Consumo de API de modelo de análisis de sentimiento basado en aspectos			
		Prueba de desempeño de consumo	Desarrollador de Backend e integración	8	4
		Creación de método de consumo	Desarrollador de Backend e integración	8	4
		Orquestación de proceso de información			
		Uso de consumo de API de voz a texto, conversión a formato TF-Record y API de modelo de análisis	Desarrollador de Backend e integración	64	4
		Manejo de excepciones	Desarrollador de Backend e integración	32	4
		Control de analíticas de validación			
		Desarrollo de modelo analítico			
		Desarrollo tabla de hechos de análisis de sentimiento de clientes y usuarios	Desarrollador de Business Intelligence	16	3
		Desarrollo tabla de hechos de procesamiento de información de grabaciones	Desarrollador de Business Intelligence	16	3
		Desarrollo tabla de hechos de indicadores de abandono, tiempo de espera y tiempo de servicio	Desarrollador de Business Intelligence	16	3
		Desarrollo tabla de hechos de análisis de grabación	Desarrollador de Business Intelligence	16	3
		Desarrollo dimensión de tiempo	Desarrollador de Business Intelligence	8	3
		Desarrollo dimensión de servicio	Desarrollador de Business Intelligence	8	3
		Desarrollo dimensión de Regional	Desarrollador de Business Intelligence	8	3
		Desarrollo dimensión de aspectos	Desarrollador de Business Intelligence	8	3
		Desarrollo dimensión de clientes	Desarrollador de Business Intelligence	8	3
		Desarrollo dimensión de asesores	Desarrollador de Business Intelligence	8	3
		Control procesamiento de información de grabaciones			
		Desarrollo de vista de procesamiento de información de grabaciones			
		Desarrollo de vista de estados de nivel 1 por servicio y regional	Desarrollador de Frontend	16	3
		Desarrollo de vista de estados de nivel 2 por registro con vista de excepciones	Desarrollador de Frontend	16	3
		Workflow de gestión			
		Registro de tickets			
		Desarrollo registro de tickets por agrupaciones	Desarrollador de Frontend	16	3
		Desarrollo registro de tickets por atención a cliente	Desarrollador de Frontend	16	3

Módulo / Fase	Historia de usuario	Tarea	Responsable	Horas	Sprint
		Desarrollo registro de <i>tickets</i> de diferencia de análisis manual y resultado de modelo	Desarrollador de <i>Frontend</i>	16	3
		Control de roles			
		Control de segregación por Regional y servicio			
		Desarrollo de segregación por Regional	Desarrollador de <i>Frontend</i>	16	3
		Desarrollo de segregación por servicio	Desarrollador de <i>Frontend</i>	16	3
		Control de segregación por rol de analista			
		Desarrollo de segregación por suma de roles por tipo de analista	Desarrollador de <i>Frontend</i>	16	3
		Desarrollo de analíticas			
		Analíticas de validación nivel 1			
		Desarrollo de filtros	Desarrollador de <i>Frontend</i>	16	4
		Desarrollo de vista de indicadores convencionales (abandono, tiempos, ANSs)	Desarrollador de <i>Frontend</i>	16	4
		Desarrollo de vista de indicadores de aspectos de análisis de sentimiento	Desarrollador de <i>Frontend</i>	32	4
		Desarrollo de vista de deltas con periodo anterior	Desarrollador de <i>Frontend</i>	32	4
		Analíticas de validación nivel 2			
		Desarrollo de vista detallada por registro	Desarrollador de <i>Frontend</i>	16	4
		Desarrollo de vista de grabación por tiempos, flujo de sentimiento de asesor y palabras claves de aspectos	Desarrollador de <i>Frontend</i>	32	4
		Pruebas integrales			
		Pruebas de modelo de análisis de sentimiento basado en aspectos	Product Owner / Analista de pruebas	16	5
		Pruebas integrales de sistema de procesamiento de información de grabaciones	Product Owner / Analista de pruebas	16	5
		Pruebas integrales de analíticas de validación	Product Owner / Analista de pruebas	48	5
		Pruebas de <i>workflow</i> de gestión	Product Owner / Analista de pruebas	16	5
		Pruebas de control de roles	Product Owner / Analista de pruebas	16	5
		Validación de usabilidad			
		Ejecución encuestas	Product Owner	6	5
		Elaboración de plan de acción	Scrum master	4	5

Fuente: Elaboración propia.

En total se planifica ejecutar el desarrollo en 5 *sprints*, cada uno de dos semanas, con un probable esfuerzo adicional derivado de la validación de usabilidad.

Por cada rol se consideran los siguientes esfuerzos:

Tabla 4. Esfuerzo por rol

Roles	Horas
<i>Scrum master</i>	400
Desarrollador de <i>Frontend</i>	272
Desarrollador de <i>Backend</i> e Integración	268
Ingeniero de datos	224
Analista de pruebas	192
<i>Product Owner</i>	177
Desarrollador de <i>Business Intelligence</i>	112
Diseñador solución <i>Business Intelligence</i>	40
Total	1.685

Fuente: Elaboración propia.

4.4.4. Presupuesto

Licencias de Plataformas de computación en la nube

Se calcula el uso de la *API Speech to Text* para 20.000 minutos en un mes:

Figura 28. Calculadora de precios de Google Cloud Platform para la API de Speech to Text (Google, 2020a)

Se calcula el uso del servicio *AI (Inteligencia artificial) Platform* para exponer el modelo de análisis de sentimiento basado en aspectos como servicio:

Figura 29. Calculadora de precios de AI Platform que permite la publicación del modelo de análisis de sentimiento basado en aspectos (Google, 2020a)

Para la plataforma Azure se tienen los siguientes costos:

- *Microsoft 365*: 20 USD por usuario por mes
- *Power Apps*: 10 USD por usuario por mes
- *Power BI Pro*: 9.99 USD por usuario por mes
- Base de datos Azure: de 0.7 a 1 USD la hora dependiendo del licenciamiento (Microsoft, 2020a) = 520 a 744 USD
- Servidor Windows para conversión a formato *TF-Record*: USD 152.57 por mes

Figura 30. Calculadora de precios de Azure para máquina virtual con licencia de Windows (Microsoft, 2020b)

Para las licencias de ambiente productivo se tiene un costo mensual de USD 3.096,98.

Automatización de tareas de validación por medio de procesamiento de lenguaje natural

4.4.5. Análisis de riesgo

Identificación de aspectos

Los aspectos para la medición de sentimiento deben ser identificados correctamente para cada negocio en particular para que los análisis se relacionen directamente con la estrategia. Por esta razón se plantean grupos focales para el levantamiento de requisitos de aspectos con diferentes roles. En todo el ciclo de proyecto se debe monitorear que los aspectos levantados al inicio del proyecto sean bien identificados y que los datos para sintonía fina sean adecuados para que el modelo maneje los aspectos estratégicos para la organización.

Obtención y limpieza de datos

En general los proyectos de análisis de información tienen sus esfuerzos más representativos en esta fase. El proyecto debe apoyarse con las áreas de tecnología y procesos del cliente para apoyar esta labor.

Precisión del modelo

La precisión del modelo depende del manejo de los anteriores dos riesgos.

Resistencia al cambio

Con la propuesta se propone de manera complementaria cambiar el proceso común de empresas de consolidación de indicadores y validación de información. Es importante apoyar el proyecto con agentes de cambio que entiendan las ventajas del nuevo proceso y mostrar entregas tempranas.

4.4.6. Lecciones aprendidas

Tal como lo plantean las técnicas ágiles, el final de cada *sprint* se debe agendar una reunión de retrospectiva para tratar los temas de mejora continua del trabajo del equipo y el cumplimiento de las metas.

4.5. Modelado del diseño de software

4.5.1. Casos de uso

El siguiente diagrama muestra los casos de uso del sistema. Se muestran los roles que interactúan con el sistema y las acciones que pueden ejecutar.



Figura 31. Casos de uso (Elaboración propia)

4.5.2. Requisitos funcionales y no funcionales

La siguiente tabla muestra los requisitos funcionales y no funcionales del sistema:

Tabla 5. Requisitos funcionales y no funcionales.

ID	Descripción
RF01	El sistema debe pasar la voz a texto de las grabaciones de atención a clientes identificando los roles de cliente y asesor manteniendo la relación con el id del caso. En caso de error debe almacenar y mostrar las excepciones.
RF02	El sistema debe codificar los textos de las grabaciones de atenciones a clientes en formato <i>TF - Record</i> . En caso de error debe almacenar y mostrar las excepciones.
RF03	El sistema debe ejecutar el modelo de análisis de sentimiento basado en aspectos en los textos de grabaciones de atenciones a clientes. En caso de error debe almacenar y mostrar las excepciones.
RF04	El sistema debe almacenar los análisis de sentimiento basado en aspectos relacionados con los casos. En caso de error debe almacenar y mostrar las excepciones.
RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes y asesores.
RF06	El sistema debe permitir registrar requisitos de gestión por agregación, es decir, en un filtro de análisis debe permitir registrar un requisito de gestión que aplique a una combinación de parámetros de entrada fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes y asesores. El requisito de gestión aplica a los registros filtrados, no a un registro en particular. El objetivo es gestionar acciones de mejora que apliquen a un comportamiento masivo, o al menos grupal. El sistema debe entregar un número de <i>ticket</i> al asesor que lo registra.
RF07	El sistema debe permitir registrar requisitos de gestión por registro. El sistema debe entregar un número de <i>ticket</i> al asesor que lo registra.
RF08	El sistema debe permitir diferencias de análisis manual con resultado de modelo. El sistema debe entregar un número de <i>ticket</i> al asesor que lo registra.
RF09	El sistema debe permitir registrar gestiones realizadas sobre los <i>tickets</i> identificados en las funcionalidades descritas en los requisitos 6, 7 y 8.

ID	Descripción
RF10	El sistema debe permitir hacer búsquedas de gestiones por los filtros fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes y asesores con el objetivo de hacer análisis de comportamientos similares o complementarios.
RF11	El sistema debe permitir visualizar indicadores de los procesos de consolidación de información, validación de información, acciones de gestión y sintonía fina de modelo a los roles que tengan el permiso correspondiente.
RNF01	El sistema debe permitir configurar roles para segregar funciones.
RNF02	El sistema debe permitir la funcionalidad y usabilidad definida en las interfaces de usuario con el objetivo de maximizar la facilidad de uso.
RNF03	A nivel de performance el sistema debe procesar la información de paso de voz a texto y análisis de sentimiento basado en aspectos de acuerdo al presupuesto disponible para el uso de <i>TPUs</i> . Se espera que las grabaciones mínimo sean procesadas en el mismo día calendario.
RNF04	Con el objetivo de mantener altos estándares de seguridad, en los servicios de <i>cloud</i> se requiere que para el acceso a datos se use <i>SSL</i> con protocolo <i>TLS</i> 1.2 para transporte y usar cifrado de datos en reposo.
RNF05	Se requiere alta disponibilidad para los servicios de AWS y Azure para mantener disponibilidad de 99.995% y un <i>RTO</i> de máximo 2 minutos.
RNF06	El tiempo de aprendizaje de uso del sistema debe ser de máximo 2 horas en promedio.
RNF07	La tasa de errores cometidos por el usuario deberá ser menor del 1% de las transacciones totales ejecutadas en el sistema.

Fuente: Elaboración propia.

4.5.3. Diagrama de actividad de validación, gestión y sintonía fina

El siguiente diagrama muestra la integración de los roles de validación de información, gestión de análisis y sintonía fina de modelo. Son los roles operativos que interactúan con el sistema.

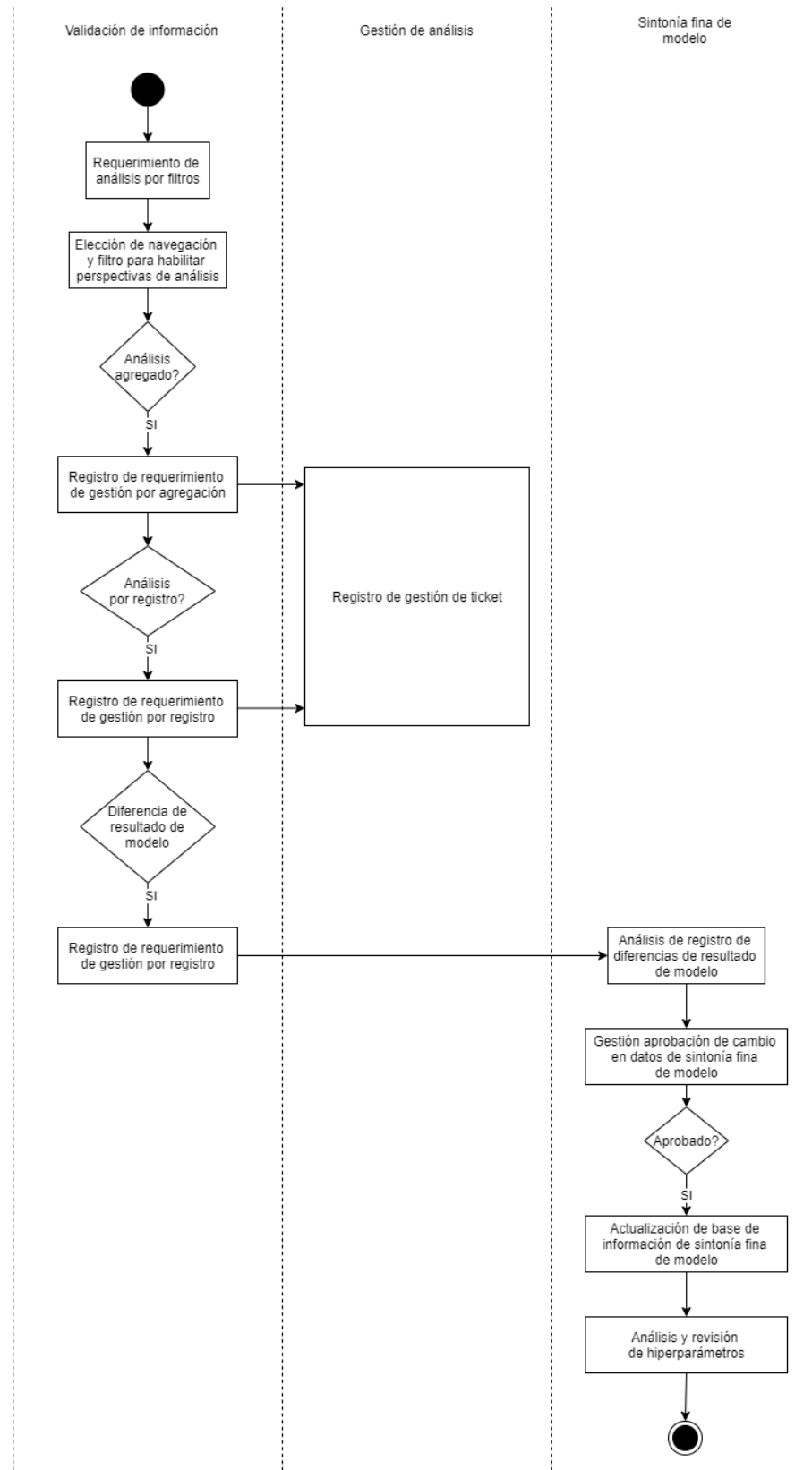


Figura 32. Diagrama de actividades de proceso de validación, gestión y sintonía fina frecuente de modelo (Elaboración propia)

4.5.4. Diagrama de estados de *tickets* de gestión

Con este diagrama se complementa la definición del diagrama de actividades del anterior numeral. Se muestra el flujo de estados de un *ticket* de gestión de información, manejado por el rol de analista de gestión:

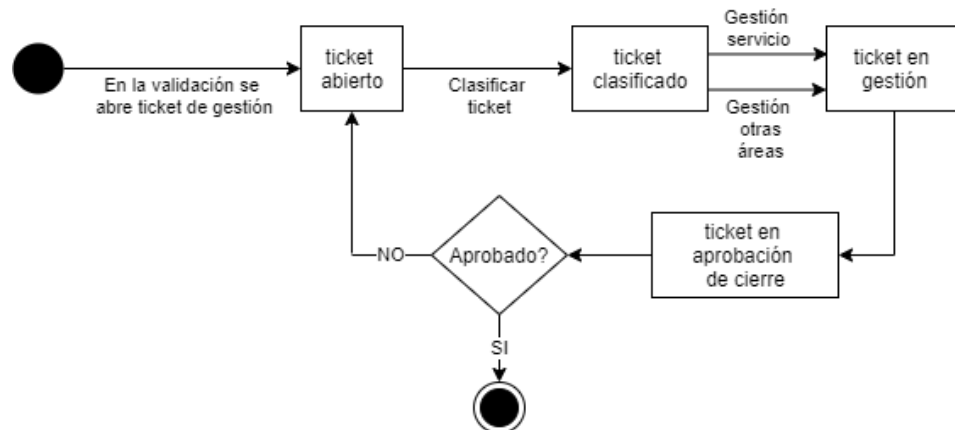


Figura 33. *Diagrama de estados para tickets de gestión de análisis* (Elaboración propia)

4.5.5. Diagrama de colaboración de procesamiento de datos

Este diagrama muestra el procesamiento de información desde la grabación de la llamada de atención al cliente, hasta obtener y relacionar el resultado del análisis de sentimiento basado en aspectos:

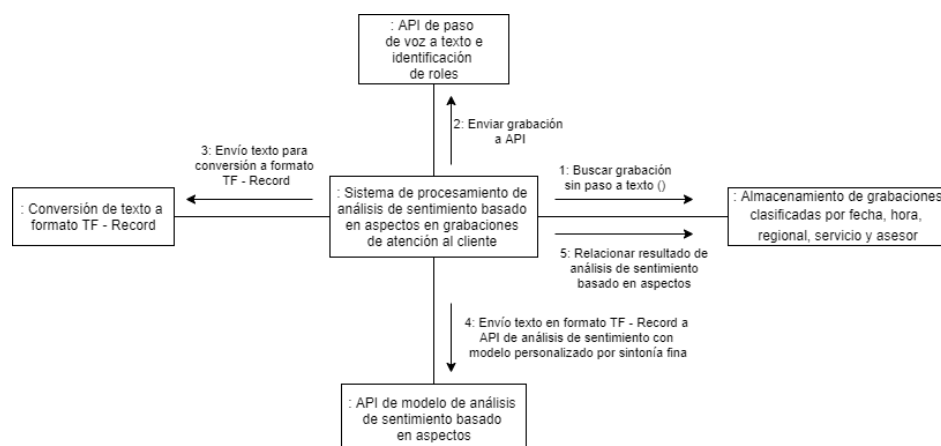


Figura 34. *Diagrama de colaboración de procesamiento de información* (Elaboración propia)

4.5.6. Diagrama de componentes de sistema

El siguiente diagrama de componentes los divide por presentación, lógica de negocio y *Backend*:

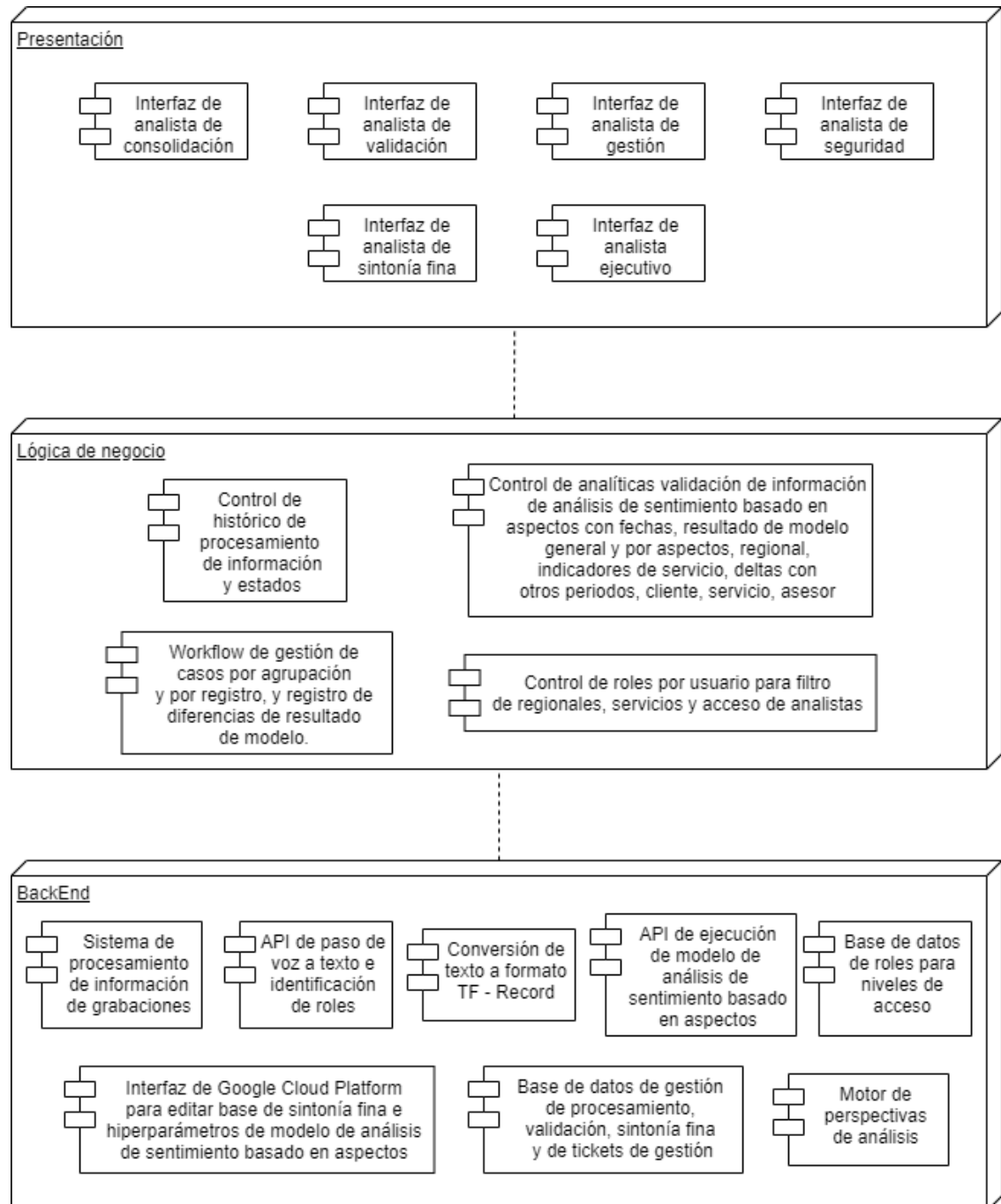


Figura 35. Diagrama de componentes de presentación, lógica de negocio y backend (Elaboración propia)

4.6. Modelo de proceso de desarrollo

Se propone el uso de *Power Apps* para consolidar una aplicación de manera incremental facilitando el uso de técnicas ágiles (Microsoft, 2020c). Este *framework* facilita la elaboración de la aplicación, se complementa con *Power BI* para el uso de analíticas y permite usar funcionalidades avanzadas personalizadas. El modelo de desarrollo parte de los prototipos, casos de uso, requisitos funcionales, diagramas de actividad, secuencia, estados y colaboración para el desarrollo. Los diagramas *UML* (lenguaje unificado de modelado) realizados son esquemáticos, no están orientados a desarrollo por modelos.

4.7. Arquitectura de solución

La solución integral se compone de los servicios de sintonía fina, uso de modelo *BERT* y *API Speech to Text* de *Google Cloud Platform*, y los servicios de base de datos, servidor, *Power BI* y *Power Apps* en *Microsoft Azure*:

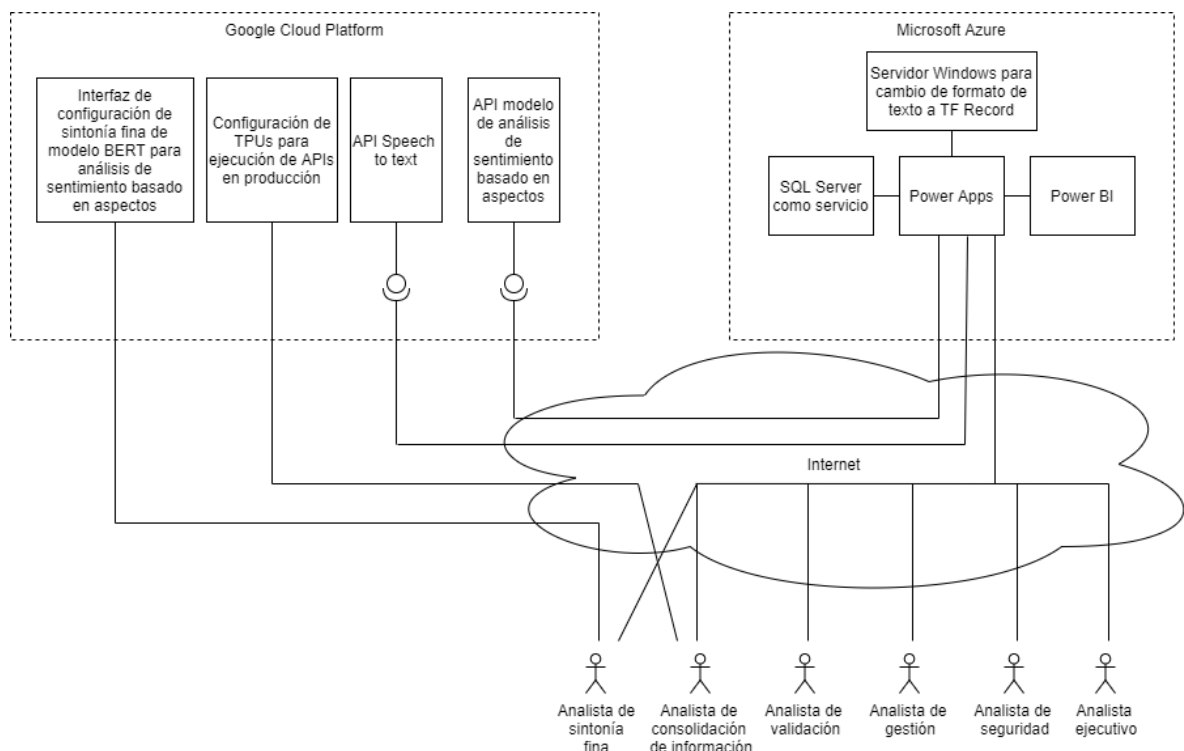


Figura 36. *Diagrama de arquitectura* (Elaboración propia)

4.8. Interfaz de usuario

Para el proceso de consolidación de procesamiento se tienen las siguientes interfaces de usuario:



Figura 37. Interfaz inicial de proceso de consolidación de información (Elaboración propia)

Esta pantalla muestra a la izquierda los menús de la aplicación, y a la derecha la pantalla inicial de consolidación de análisis. El usuario puede seleccionar fecha inicio y fecha fin, y puede seleccionar el registro de un servicio y una regional para ver el detalle de los registros en la siguiente pantalla.

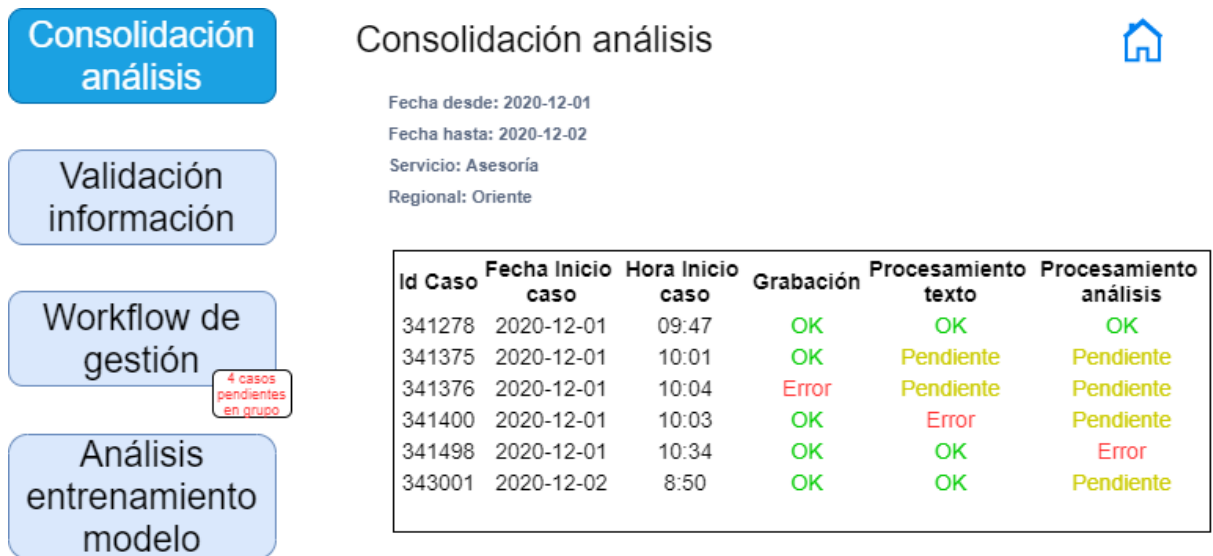


Figura 38. Interfaz proceso de consolidación de información de registro de casos (Elaboración propia)

En esta pantalla se muestra el detalle de los registros y el avance en el almacenamiento de la grabación, el paso de voz a texto en formato *TF-Record* y procesamiento de análisis. Se detalla el estado por registro. Al dar clic en algún registro se mostrará el detalle del procesamiento, y en caso de error, dará opciones de reintento en la siguiente pantalla.

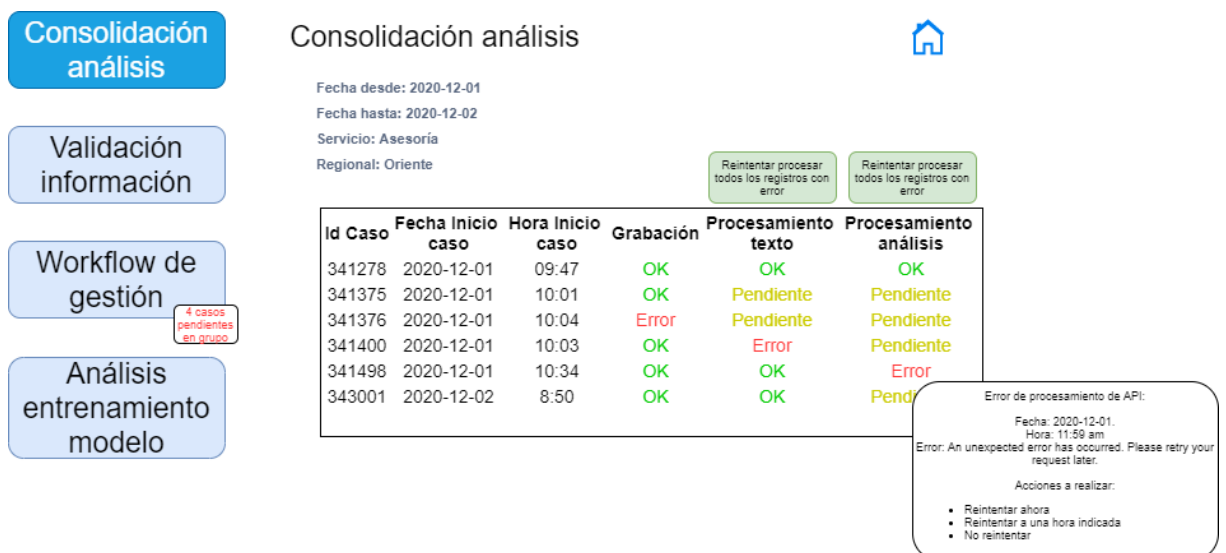


Figura 39. Interfaz proceso de consolidación de información con opciones de acciones sobre registros detallados (Elaboración propia)

Al dar clic en los registros de grabaciones en cada proceso se mostrará un popup con fecha y hora de ejecución, y si hay error se muestra el error detallado, teniendo en cuenta que es un usuario técnico el que va a ser uso de esta opción. Aunque el sistema cuenta con un proceso de reintentos, se da la opción de iniciar un reintento manual con el análisis del usuario. En la parte superior de los campos se da opción de reintentar todos los registros en error filtrados en el reporte.

Para el proceso de validación de información se tienen las siguientes interfaces:



Figura 40. Interfaz inicial de proceso de validación de información (Elaboración propia)

En esta pantalla se muestra una perspectiva general del análisis de sentimiento basado en aspectos para el día actual, el usuario puede cambiar el filtro para ver los datos de un día o rango de fechas diferente. Se permite filtrar por Regional, Servicio, Aspectos, Segmento de clientes, Área de asesor y Asesor. Se muestran indicadores convencionales de servicio, medición de sentimiento global y por aspecto, se muestra por regional el sentimiento global y la tendencia con respecto al día anterior comparado con el indicador de cantidad de asesores por número de contactos. Se muestra una nube de palabras en las que se ven las más usadas por el cliente marcadas con un color que muestra el sentimiento del cliente. Desde esta pantalla se puede seleccionar una regional, un aspecto combinado con la calificación y una palabra para desplegar la siguiente pantalla.



Figura 41. Interfaz que muestra los registros detallados de los casos (Elaboración propia)

En esta pantalla se muestra la información de los registros detallados filtrados en la pantalla anterior. Se puede desplegar u ocultar la información de caso, cliente, asesor y el análisis de aspectos. Al seleccionar un registro, se despliega la siguiente pantalla.

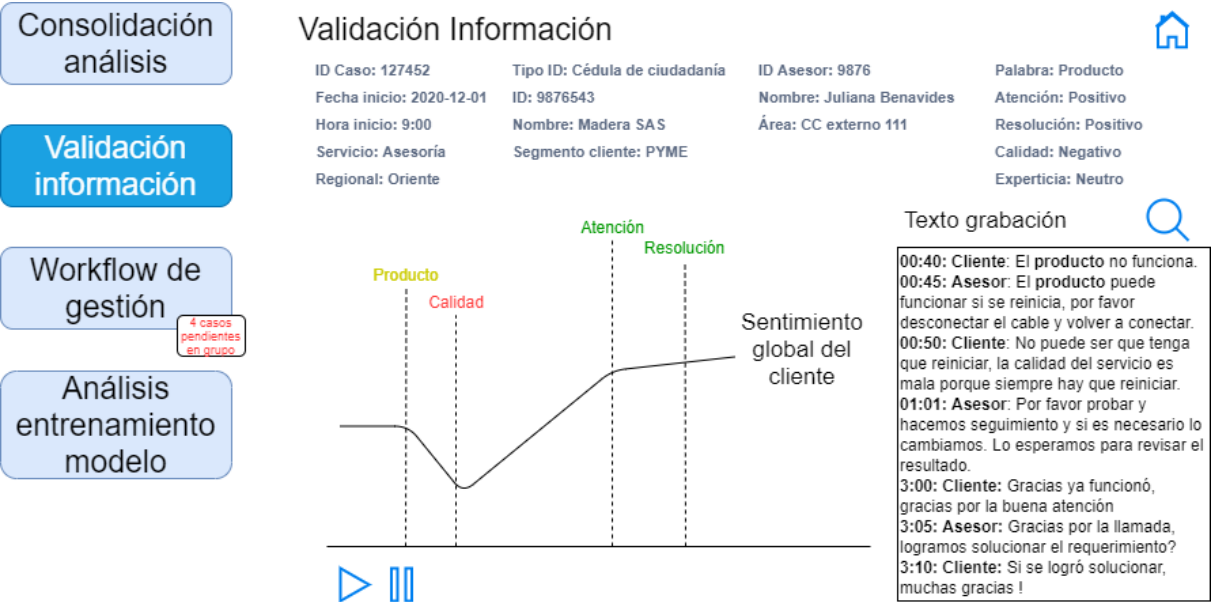


Figura 42. Interfaz de reproducción de grabación (Elaboración propia)

En esta pantalla se permite reproducir la grabación, visualizando el texto. Al reproducir la grabación se va marcando la sección de texto que corresponde. El texto permite buscar palabras. Se muestra las secciones de la llamada en donde se menciona la palabra filtrada, y las calificaciones de aspectos marcadas por colores, dependiendo de la calificación.

4.9. Técnicas, lenguajes y herramientas de implementación

El proyecto va a ser implementado usando técnicas ágiles. En el plan se proponen las actividades con base en *sprints* de 2 semanas. En el desarrollo de los *sprints* se manejan las reuniones propuestas por *Scrum*, la reunión diaria, planificación, revisión y retrospectiva. Aunque para el plan se proponen las estimaciones, con el objetivo de tener una perspectiva de tiempo para las entregas, estas estimaciones se realizan con el equipo de desarrollo con la aplicación *Scrum Poker Cards*. La unidad de estimaciones se lleva en horas. El control del tiempo tomado en las tareas y las asignaciones se manejan en JIRA. De JIRA se consulta el *Burndown Chart* para llevar el control diario del avance y cumplimiento.

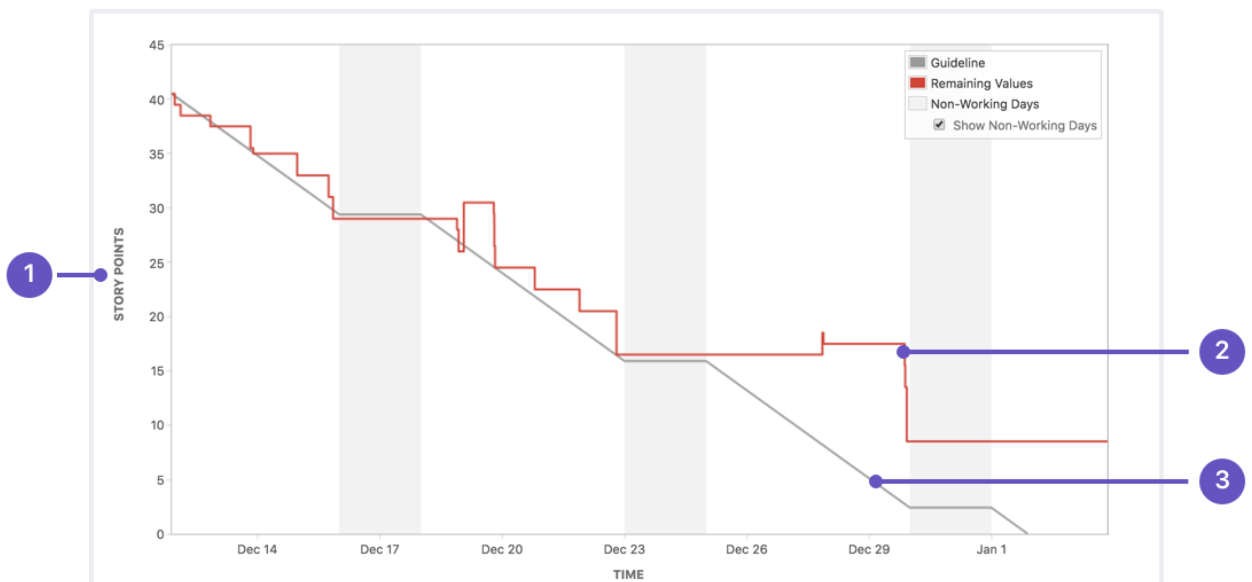


Figura 43. *Burndown Chart de JIRA* (Rehkopf, 2020)

1. En el eje vertical están los puntos de la historia. En el caso del proyecto la unidad correspondiente a los puntos son horas.
2. En la gráfica en rojo se ve el avance real consultado en las reuniones diarias.
3. En la gráfica gris se muestra el avance ideal para terminar el *sprint* cumpliendo los objetivos.

En la siguiente tabla se muestran los tipos de desarrollo con las herramientas y lenguajes usados.

Tabla 6. Tipos de desarrollos, herramientas y lenguajes de programación

Tipos de desarrollos del proyecto	Herramientas	Lenguajes
Interfaz de usuario	<i>Power Apps</i> <i>Power BI</i>	<i>.NET</i> <i>T-SQL</i>
Control de roles de seguridad	<i>Power Apps</i>	<i>.NET</i> <i>T-SQL</i>
<i>Workflow</i> de tickets de gestión	<i>Power Apps</i>	<i>.NET</i> <i>T-SQL</i>
Consolidación información de análisis	<i>Azure SQL Server service</i> <i>SSIS - SQL Server Integration Services</i>	<i>T-SQL</i>
Procesamiento y Transformación de información antes de consumo de APIs	<i>.NET Framework</i>	<i>.NET</i> <i>T-SQL</i>
Consumo de APIs <i>Speech to Text</i> y <i>AI Platform</i>	<i>GCP API Speech to Text</i> <i>GCP AI Platform</i>	<i>.NET</i>
Sintonía fina de modelo <i>BERT</i> pre-entrenado	<i>GCP (Google Cloud Platform) Command Line</i> <i>GCP Storage Bucket</i> <i>GCP Cloud TPU</i> <i>GCP Compute Engine VM</i> <i>Github</i>	<i>Python</i> <i>Tensorflow</i>

Fuente: Elaboración propia.

4.10. Tareas de sintonía fina de modelo

Las tareas de sintonía fina no hacen parte del sistema de gestión. Son acciones que se ejecutan en la implementación y en la sintonía fina sobre el modelo configurado en *Google Cloud Platform*. El sistema interactúa con el modelo para ejecutar los análisis, pero no tiene la funcionalidad de cambiar los parámetros de sintonía fina, estos cambios se deben hacer directamente en la interfaz de *Google Cloud Platform*.

4.10.1 Definición de Requisitos

- Definición de aspectos o entidades de análisis
- Definición de datos de sintonía fina, validación y producción
- Definición de indicadores de análisis de sentimiento basado en aspectos

4.10.2 Análisis y diseño

- Análisis y diseño de formato y limpieza de datos de sintonía fina, validación y productivos
- Exploración, análisis y perfilamiento de datos y preprocesamiento
- Análisis y diseño de indicadores de análisis de sentimiento basado en aspectos

4.10.3 Implementación

- Codificación de datos
- Configuración de modelo pre-entrenado
- Crear cargas de datos
- Configuración de Optimizador y Agendamiento
- Definición de métricas de desempeño de modelo
- Creación de ciclo de sintonía fina
- Cargando y evaluando modelo
- Implementación de indicadores de análisis de sentimiento basado en aspectos

4.11. Pruebas y validación

4.11.1. Diseño de pruebas

El diseño de las pruebas se realizó tomando los elementos de diseño requisitos funcionales y no funcionales, casos de uso, diagrama de componentes.

Para cada caso se analizaron los escenarios de las funcionalidades para establecer los roles, resultado esperado y precondiciones para que la prueba sea válida. Se espera que en la ejecución de pruebas el analista diligencie si el resultado es exitoso y las observaciones para compartirlo con el equipo de desarrollo en caso de requerir ajustes.

Tabla 7. Casos de prueba

Nro	Req.	Descripción requisite	Caso de Uso	Tipo de Caso	Caso de prueba	Roles o actores	Pre-condición	Resultado esperado	Resultado Exitoso	Observaciones
1	RF01	El sistema debe pasar la voz a texto de las grabaciones de atención a clientes identificando los roles de cliente y asesor manteniendo la relación con el <i>id</i> del caso. En caso de error debe almacenar y mostrar las excepciones.	El sistema debe pasar la voz de grabaciones de atención a clientes a texto	Funcional	El sistema pasa la voz de la grabación a texto	Analista de consolidación de información	La grabación debe estar almacenada con los datos de fecha, hora, regional, servicio, asesor, cliente	Debe realizar el paso de voz a texto		
2	RF01	El sistema debe pasar la voz a texto de las grabaciones de atención a clientes identificando los roles de cliente y asesor manteniendo la relación con el <i>id</i> del caso. En caso de error debe almacenar y mostrar las excepciones.	El sistema debe identificar los roles de cliente y asesor en la llamada	Funcional	El sistema debe identificar los roles de cliente y asesor	Analista de consolidación de información	La grabación debe identificar las secciones en las que habla el cliente y el asesor	Debe realizar la identificación de los roles de cliente y asesor para que el texto corresponda a la información de la llamada		
3	RF02	El sistema debe codificar los textos de las grabaciones de atenciones a clientes en formato <i>TF - Record</i> . En caso de error debe almacenar y mostrar las excepciones.	El sistema debe codificar los textos de las grabaciones en formato <i>TF - Record</i>	Funcional	El sistema debe codificar los textos de las grabaciones en formato <i>TF - Record</i>	Analista de consolidación de información	La grabación de la atención telefónica debe estar convertida a texto	Debe realizar la conversión del texto de la grabación a formato <i>TF - Record</i> manteniendo la identificación de roles		

Nro	Req.	Descripción requisite	Caso de Uso	Tipo de Caso	Caso de prueba	Roles o actores	Pre-condición	Resultado esperado	Resultado Exitoso	Observaciones
4	RF03	El sistema debe ejecutar el modelo de análisis de sentimiento basado en aspectos en los textos de grabaciones de atenciones a clientes. En caso de error debe almacenar y mostrar las excepciones.	El sistema debe ejecutar el modelo de análisis de sentimiento en los datos codificados en formato <i>TF - Record</i> . La información se debe complementar con los aspectos evaluados	Funcional	El sistema debe ejecutar el modelo de análisis de sentimiento en los datos <i>TF-Record</i> y debe marcar los aspectos evaluados	Analista de consolidación de información	La información de grabaciones debe estar almacenada en formato <i>TF-Record</i>	Debe ejecutar el modelo en los datos en formato <i>TF - Record</i> y debe remarcar los aspectos evaluados		
5	RF04	El sistema debe almacenar los análisis de sentimiento basado en aspectos relacionados con los casos. En caso de error debe almacenar y mostrar las excepciones.	El sistema debe relacionar los resultados del modelo de análisis de sentimiento basado en aspectos con las grabaciones de atención al cliente	Funcional	El sistema debe relacionar los resultados del modelo de análisis de sentimiento basado en aspectos con las grabaciones de atención al cliente	Analista de consolidación de información	Se debe haber ejecutado previamente el modelo de análisis de sentimiento basado en aspectos sobre los datos en <i>TF-Record</i> de las grabaciones	Debe relacionarse la información de los resultados del modelo de análisis de sentimiento basado en aspectos con la información del caso		
6	RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes, asesores y clientes.	El sistema debe permitir analizar la información por fecha y hora	Funcional	El sistema debe permitir analizar la información por fecha y hora	Analista de validación de información, Analista ejecutivo	Se debe haber cumplido con el proceso de procesamiento y relacionamiento de información	Debe permitir filtrar la información por fecha y hora		
7	RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes, asesores y clientes.	El sistema debe permitir analizar la información por regional	Funcional	El sistema debe permitir analizar la información por regional	Analista de validación de información, Analista ejecutivo	Se debe haber cumplido con el proceso de procesamiento y relacionamiento de información	Debe permitir filtrar la información por Regional		

Nro	Req.	Descripción requisito	Caso de Uso	Tipo de Caso	Caso de prueba	Roles o actores	Pre-condición	Resultado esperado	Resultado Exitoso	Observaciones
8	RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes, asesores y clientes.	El sistema debe permitir analizar la información por servicio	Funcional	El sistema debe permitir analizar la información por servicio	Analista de validación de información, Analista ejecutivo	Se debe haber cumplido con el proceso de procesamiento y relacionamiento de información	Debe permitir filtrar la información por Servicio		
9	RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento, asesores y clientes.	El sistema debe permitir analizar la información por aspectos	Funcional	El sistema debe permitir analizar la información por aspectos	Analista de validación de información, Analista ejecutivo	Se debe haber cumplido con el proceso de procesamiento y relacionamiento de información	Debe permitir filtrar la información por Aspectos		
10	RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes, asesores y clientes.	El sistema debe permitir analizar la información por rangos de evaluación de sentimiento	Funcional	El sistema debe permitir analizar la información por rangos de evaluación de sentimiento	Analista de validación de información, Analista ejecutivo	Se debe haber cumplido con el proceso de procesamiento y relacionamiento de información	Debe permitir filtrar la información por Rangos de evaluación de sentimiento		
11	RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes, asesores y clientes.	El sistema debe permitir analizar la información por Asesores	Funcional	El sistema debe permitir analizar la información por Asesores	Analista de validación de información, Analista ejecutivo	Se debe haber cumplido con el proceso de procesamiento y relacionamiento de información	Debe permitir filtrar la información por Asesores		

Nro	Req.	Descripción requisite	Caso de Uso	Tipo de Caso	Caso de prueba	Roles o actores	Pre-condición	Resultado esperado	Resultado Exitoso	Observaciones
12	RF05	El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes, asesores y clientes.	El sistema debe permitir analizar la información por Clientes	Funcional	El sistema debe permitir analizar la información por Clientes	Analista de validación de información, Analista ejecutivo	Se debe haber cumplido con el proceso de procesamiento y relacionamiento de información	Debe permitir filtrar la información por Clientes		
13	RF06	El sistema debe permitir registrar requisitos de gestión por agregación, es decir, en un filtro de análisis debe permitir registrar un requisito de gestión que aplique a una combinación de parámetros de entrada fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes y asesores. El requisito de gestión aplica a los registros filtrados, no a un registro en particular. El objetivo es gestionar acciones de mejora que apliquen a un comportamiento masivo, o al menos grupal. El sistema debe entregar un número de <i>ticket</i> al asesor que lo registra.	El sistema debe permitir registrar requisitos de gestión sobre filtros realizados especificando la observación del análisis	Funcional	El sistema debe permitir registrar requisitos de gestión sobre filtros realizados especificando la observación del análisis	Analista de validación de información, Analista ejecutivo	El registro de requisito de gestión parte del análisis de validación de información. Debe ejecutarse el análisis para identificar los requisitos de gestión	Debe permitir registrar requisitos de gestión sobre análisis ejecutados. En el registro se debe identificar el filtro ejecutado y la observación del analista de validación		

Nro	Req.	Descripción requisito	Caso de Uso	Tipo de Caso	Caso de prueba	Roles o actores	Pre-condición	Resultado esperado	Resultado Exitoso	Observaciones
14	RF07	El sistema debe permitir registrar requisitos de gestión por registro. El sistema debe entregar un número de <i>ticket</i> al asesor que lo registra.	El sistema debe permitir registrar requisitos de gestión sobre grabaciones de clientes específicas	Funcional	El sistema debe permitir registrar requisitos de gestión sobre grabaciones de clientes específicas	Analista de validación de información, Analista ejecutivo	El registro de requisito de gestión parte del análisis de validación de información. Debe ejecutarse el análisis para identificar los requisitos de gestión	Debe permitir registrar requisitos de gestión sobre registros de atenciones específicas de clientes		
15	RF08	El sistema debe permitir registrar diferencias de análisis manual con resultado de modelo. El sistema debe entregar un número de <i>ticket</i> al asesor que lo registra.	El sistema debe permitir registrar diferencias de análisis manual con resultado del modelo de análisis de sentimiento basado en aspectos	Funcional	El sistema debe permitir registrar diferencias de análisis manual con resultado del modelo de análisis de sentimiento basado en aspectos	Analista de validación de información, Analista ejecutivo	El registro de requisito de diferencia de resultados de modelo parte del análisis de validación de información. Debe ejecutarse el análisis para identificar los requisitos de diferencias	Debe permitir registrar requisitos de diferencias con resultado de modelo sobre grabaciones específicas de atención de clientes		
16	RF09	El sistema debe permitir registrar gestiones realizadas sobre los <i>tickets</i> identificados en las funcionalidades descritas en los requisitos 6, 7 y 8.	El sistema debe permitir registrar cambios de estado de <i>tickets</i> de gestión y observaciones de las gestiones realizadas	Funcional	El sistema debe permitir registrar cambios de estado de <i>tickets</i> de gestión y observaciones de las gestiones realizadas	Analista de gestión, Analista de sintonía fina	Para iniciar la gestión y cambios de estado se deben haber ingresado previamente los <i>tickets</i> por en el proceso de validación	Debe permitir ingresar comentarios de gestión y cambios de estados de <i>tickets</i>		
17	RF10	El sistema debe permitir registrar gestiones realizadas sobre los <i>tickets</i> identificados en las funcionalidades descritas en los requisitos 6, 7 y 8.	El sistema debe permitir consultar el histórico de gestión de <i>tickets</i> por filtros para analizar el comportamiento histórico de validaciones y gestiones.	Funcional	El sistema debe permitir consultar el histórico de gestión de <i>tickets</i> por filtros para analizar el comportamiento histórico de validaciones y gestiones.	Analista de gestión, Analista de sintonía fina	Para consultar el histórico debe existir gestión de <i>tickets</i>	Debe permitir consultar el histórico de gestiones		
18	RF11	El sistema debe permitir visualizar indicadores de los procesos de consolidación de información, validación de información, acciones de	El sistema debe permitir visualizar indicadores de los procesos de consolidación de información, validación de información, acciones de	Funcional	El sistema debe permitir visualizar indicadores de los procesos de consolidación de información, validación de información,	Analista ejecutivo	Para consultar los indicadores de los procesos deben haberse ejecutado acciones en cada uno previamente	Debe permitir visualizar indicadores de los procesos de consolidación de información, validación de información, acciones de		

Nro	Req.	Descripción requisite	Caso de Uso	Tipo de Caso	Caso de prueba	Roles o actores	Pre-condición	Resultado esperado	Resultado Exitoso	Observaciones
		gestión y sintonía fina de modelo a los roles que tengan el permiso correspondiente.	gestión y sintonía fina de modelo		acciones de gestión y sintonía fina de modelo			gestión y sintonía fina de modelo		
19	RNF 01	El sistema debe permitir configurar roles para segregar funciones.	El sistema debe permitir segregar permisos por regional, servicio y roles de analistas	No funcional	El sistema debe permitir segregar permisos por regional, servicio y roles de analistas	Analista ejecutivo	-	Debe permitir definir los roles de cada analista que interviene en el sistema por regional, servicio y roles de analistas		
20	RNF 02	El sistema debe permitir la funcionalidad y usabilidad definida en los prototipos con el objetivo de maximizar la facilidad de uso.	El sistema debe permitir la funcionalidad y usabilidad definida en los prototipos con el objetivo de maximizar la facilidad de uso.	No funcional	El sistema debe permitir la funcionalidad y usabilidad definida en los prototipos con el objetivo de maximizar la facilidad de uso.	Todos	-	Debe alcanzar al menor una puntuación de 68 con el test definido por <i>System Usability Scale</i>		
21	RNF 03	A nivel de performance el sistema debe procesar la información de paso de voz a texto y análisis de sentimiento basado en aspectos de acuerdo al presupuesto disponible para el uso de <i>TPUs</i> . Se espera que las grabaciones mínimo sean procesadas en el mismo día calendario.	El sistema debe procesar el paso de voz a texto en el mismo día en que se ejecuta la atención al cliente	No funcional	El sistema debe procesar el paso de voz a texto en el mismo día en que se ejecuta la atención al cliente	Todos	-	Debe ejecutarse el proceso de paso de voz a texto en el mismo día en que ocurre la atención al cliente		
22	RNF 04	A nivel de <i>performance</i> el sistema debe procesar la información de paso de voz a texto y análisis de sentimiento basado en aspectos de acuerdo al presupuesto disponible para el uso de <i>TPUs</i> . Se espera que las	El sistema debe procesar el modelo de análisis de sentimiento basado en aspectos en el mismo día en que se ejecuta la atención al cliente	No funcional	El sistema debe procesar el modelo de análisis de sentimiento basado en aspectos en el mismo día en que se ejecuta la atención al cliente	Todos	-	Debe ejecutarse el modelo de análisis de sentimiento basado en aspectos en el mismo día en que ocurre la atención al cliente		

Nro	Req.	Descripción requisite	Caso de Uso	Tipo de Caso	Caso de prueba	Roles o actores	Pre-condición	Resultado esperado	Resultado Exitoso	Observaciones
		grabaciones mínimo sean procesadas en el mismo día calendario.								

Fuente: Elaboración propia.

4.11.2. Validación de usabilidad

Para validar la usabilidad se usará la herramienta *System Usability Scale (SUS)* (Brooke, 2020). En la metodología se propone el uso de la siguiente encuesta:

		Muy en desacuerdo				Totalmente de acuerdo
1.	Creo que usaría este Automatización de tareas de validación frecuentemente.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.	Encuentro este Automatización de tareas de validación innecesariamente complejo.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.	Creo que el Automatización de tareas de validación fue fácil de usar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.	Creo que necesitaría ayuda de una persona con conocimientos técnicos para usar este Automatización de tareas de validación.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.	Las funciones de este Automatización de tareas de validación están bien integradas.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.	Creo que el Automatización de tareas de validación es muy inconsistente.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7.	Imagino que la mayoría de la gente aprendería a usar este Automatización de tareas de validación en forma muy rápida.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8.	Encuentro que el Automatización de tareas de validación es muy difícil de usar.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9.	Me siento confiado al usar este Automatización de tareas de validación.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10.	Necesité aprender muchas cosas antes de ser capaz de usar este Automatización de tareas de validación.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figura 44. Encuesta *System Usability Scale (SUS)* (Brooke, 2020)

Los resultados en una escala de 0-40 se multiplican por 2.5 para obtener el resultado de 0-100. El objetivo de la aplicación según el estándar es obtener un puntaje igual o superior a 68.

5. Conclusiones y trabajo futuro

5.1. Conclusiones

En el presente trabajo se ha logrado planificar una solución de validación de información para analizar el sentimiento basado en aspectos de las atenciones a clientes o usuarios, cambiando la definición de las actividades operativas a analíticas. Con la planificación se cumple el objetivo de altos ejecutivos de monitorear todos los contactos de los clientes con la compañía, identificando las opiniones destacadas y habilitando gestiones sobre comportamientos masivos o individuales.

Se validó en fuentes académicas el tipo de modelo a usar para el análisis de sentimiento basado en aspectos como una tarea de procesamiento de lenguaje natural. Se revisan las opciones usadas en soluciones similares de procesamiento de lenguaje natural concluyendo que el modelo *BERT* y sus derivados tienen una mejor precisión para solucionar este tipo de problemas. El rápido avance de este tipo de tecnologías exige una investigación constante para evaluar formas novedosas para abordar este tipo de requisitos.

Se revisaron los proveedores de servicios de computación en la nube, evaluando las mejores opciones a nivel de *hardware* y *software* para la implementación del modelo. Se concluye que actualmente *Google Cloud Platform* ofrece la mejor opción de hardware y software para usar la *API* de paso de voz a texto y de análisis de sentimiento basado en aspectos.

Como parte de la planificación se incluye el desarrollo de un flujo de trabajo de análisis de información que usa los resultados del modelo para la toma de decisiones estratégicas sobre el proceso de atención a comunicaciones de usuarios, y el planteamiento de un proceso organizacional complementario al funcionamiento de la solución.

En el proceso se plantean roles que al hacer uso de la herramienta hacen parte de la transformación de las actividades y el alcance del análisis, para pasar de análisis de muestras poco precisos, a una perspectiva general de los contactos de todos los clientes para potencialmente analizar todos los canales de comunicación con el cliente, teniendo en cuenta que se analiza voz y texto.

Como parte de la planificación se incluye la toma de requisitos de clientes para la adaptación del modelo a necesidades particulares a nivel de aspectos, terminología, y flujo de trabajo. Se plantea una planificación flexible con uso de técnicas ágiles para personalizar los aspectos a analizar de determinado negocio con potencial de integrar las herramientas convencionales de TI que ya se utilizan para integrar el flujo corporativo.

5.2. Líneas de trabajo futuro

Se ha concluido que el mayor aporte del presente trabajo es plantear una planificación con un plazo relativamente corto para implementar un proceso organizacional de validación de información y una herramienta que lo apalanca, que cambia la forma en que las empresas perciben las opiniones de sus clientes por canales de atención al cliente. Sería posible abordar objetivos más ambiciosos como:

- Evaluar sentimiento de atenciones ejecutadas por *bots*.
- Minimizar la interacción humana en procesos repetitivos que puedan ser codificados.
- Complementar la base de información de entrenamiento de modelos de *AI* iterativamente con el conocimiento de expertos abordando los indicadores más relevantes. El uso de herramientas como *Dialogflow* (Google, 2020b) para automatizar la atención a clientes requiere del entrenamiento constante.
- Dar un enfoque a las estrategias de automatización de servicio al cliente orientadas en el sentir del usuario, más que en hacer laberintos de *IVR* con voces automatizadas.
- Automatizar procesos de negocio maximizando la experiencia de los clientes.

Se va a continuar investigando los temas de las tareas del *WorkFlow SemEval* y las herramientas de *cloud computing* para implementar este tipo de soluciones. Con el paso de los años los objetivos de las tareas van a ser más ambiciosos y las herramientas usadas más sofisticadas.

6. Bibliografía

- Alfaro, E. (2010). *ConocerElAutor*. Obtenido de El ABC del customer experience: <https://conoceralautor.es/ficheros/doc/00119.pdf>
- Alford, A. (2020). *InfoQ*. Obtenido de Microsoft Open-Sources ONNX Acceleration for BERT AI Model: <https://www.infoq.com/news/2020/01/microsoft-bert-acceleration/>
- Amazon. (2020). *Amazon Web Services*. Obtenido de FPGA-Accelerated Deep-Learning Inference with Binarized Neural Networks: <https://aws.amazon.com/marketplace/pp/Missing-Link-Electronics-Inc-FPGA-Accelerated-Deep/B0784D5WJK>
- Apex. (2020). *Advanced Analytics*. Obtenido de Advanced Speech Analytics: <https://apexinsights.cx/>
- Brooke, J. (2020). *Usability*. Obtenido de System Usability Scale (SUS): <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.htm>
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020a). *dccuchile*. Obtenido de dccuchile / beto: <https://github.com/dccuchile/beto>
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020b). *Google Research*. Obtenido de BETO Examples: https://colab.research.google.com/drive/1uRwg4UmPgYlqGYY4gW_Nsw9782GFJbPt
- CE Noticias Financieras English. (2019). *ProQuest*. Obtenido de How will Google BERT help you find what you're looking for easier?: <https://bv.unir.net:2257/docview/2326418941?pq-origsite=summon>
- Cloudmersive. (2020). *Cloudmersive*. Obtenido de Powerful Natural Language Processing APIs: <https://cloudmersive.com/nlp-api>
- Devlin, J., Markowitz, D., Rao, A., Pu, J., Citro, C., Châtel, G., & Asmro, A. (2020). *Github*. Obtenido de google-research / bert: <https://github.com/google-research/bert/blob/master/README.md>

- Google. (2018). *Google Research*. Obtenido de BERT End to End (Fine-tuning + Predicting) with Cloud TPU: Sentence and Sentence-Pair Classification Tasks: https://colab.research.google.com/github/tensorflow/tpu/blob/master/tools/colab/bert_finetuning_with_cloud_tpus.ipynb#scrollTo=HUBP35oCDmbF
- Google. (2019). *Google Research*. Obtenido de Welcome To Colaboratory: https://colab.research.google.com/notebooks/intro.ipynb#scrollTo=5fCEDCU_qrC0
- Google. (2020a). *Google Cloud*. Obtenido de Google Cloud Pricing Calculator: <https://cloud.google.com/products/calculator>
- Google. (2020b). *Google Cloud*. Obtenido de Dialogflow: https://cloud.google.com/dialogflow?utm_source=google&utm_medium=cpc&utm_campaign=latam-CO-all-es-dr-BKWS-all-all-trial-b-dr-1009133-LUAC0009159&utm_content=text-ad-none-any-DEV_c-CRE_436253090535-ADGP_BKWS%20%7C%20Multi%20~%20Machine%20Learning%20%7C%20D
- Google. (2020c). *Google Cloud*. Obtenido de BERT FineTuning with Cloud TPU: Sentence and Sentence-Pair Classification Tasks: <https://cloud.google.com/tpu/docs/tutorials/bert>
- Gupta, S. (2018). *ProQuest*. Obtenido de Google Open Sources its BERT Model Source Code: <https://bv.unir.net:2257/docview/2138312944?pq-origsite=summon>
- Jangamreddy, N. (2019). *Research Gate*. Obtenido de A Survey on Specialised Hardware for Machine Learning: https://www.researchgate.net/profile/Nikhil_Reddy7/publication/334647302_A_Survey_on_Specialised_Hardware_for_Machine_Learning/links/5d3842b3299bf1995b457c8d/A-Survey-on-Specialised-Hardware-for-Machine-Learning.pdf
- Jawandhiya, P. (2018). Hardware Design for Machine. *International Journal of Artificial Intelligence and Applications (IJAlA)*, Vol.9, No.1, 22.
- Jouppi, N. P., Young, C., Patil, N., Patterson, D., Agrawal, G., Bajwa, R., . . . Dau, M. (2017). *Research Gate*. Obtenido de In-Datcenter Performance Analysis of a Tensor Processing Unit: https://www.researchgate.net/profile/Norman_Jouppi/publication/319856813_In-Datcenter_Performance_Analysis_of_a_Tensor_Processing_Unit/links/599928e60f7

e9b3edb16779f/In-Datcenter-Performance-Analysis-of-a-Tensor-Processing-Unit.pdf

Lappemana, J., Clark, R., Evans, J., Sierra-Rubiac, L., & Gordon, P. (2020). *Science Direct*. Obtenido de Studying social media sentiment using human: <https://www.sciencedirect.com/science/article/pii/S2215016120300868>

Li, Q., & Ping, D. (2020). *Amazon Web Services*. Obtenido de Fine-tuning a PyTorch BERT model and deploying it with Amazon Elastic Inference on Amazon SageMaker: <https://aws.amazon.com/blogs/machine-learning/fine-tuning-a-pytorch-bert-model-and-deploying-it-with-amazon-elastic-inference-on-amazon-sagemaker/>

Li, Q., Li, S., Zhang, S., Hu, J., & Hu, J. (2019). *MDPI*. Obtenido de A Review of Text Corpus-Based Tourism Big Data Mining: <https://www.mdpi.com/2076-3417/9/16/3300>

Lixiang, L., Zhigang, G., Zhiming, W., & Liang, Y. (2019). *Alibaba Cloud*. Obtenido de Perseus-BERT: Industry-leading BERT Training Solution Optimized for Performance: https://www.alibabacloud.com/blog/perseus-bert-industry-leading-bert-training-solution-optimized-for-performance_594717

Manandhar, S., Androutsopoulos, I., Galanis, D., Papageorgiou, H., Pavlopoulos, J., & Pontiki, M. (2014a). *SemEval-2014*. Obtenido de SemEval-2014 Task 4: <http://alt.qcri.org/semeval2014/task4/data/uploads/restaurants-trial.xml>

Manandhar, S., Androutsopoulos, I., Galanis, D., Papageorgiou, H., Pavlopoulos, J., & Pontiki, M. (2014b). *SemEval-2014 Task 4*. Obtenido de SemEval-2014 : Semantic Evaluation Exercises: <http://alt.qcri.org/semeval2014/task4/>

Matarranz, A. (2017). *Youtube*. Obtenido de Cuando usar las diferentes herramientas de Analítica de Texto - MeaningCloud: <https://www.youtube.com/watch?v=6mjQ6mZJZXE>

MeaningCloud. (2020). *Meaning Cloud*. Obtenido de Análisis de Sentimiento: <https://www.meaningcloud.com/es/productos/analisis-de-sentimiento>

Microsoft. (2020a). *Microsoft Azure*. Obtenido de Azure SQL Database pricing: <https://azure.microsoft.com/en-us/pricing/details/sql-database/single/>

Microsoft. (2020b). *Microsoft Azure*. Obtenido de Windows Virtual Machines Pricing: <https://azure.microsoft.com/en-us/pricing/details/virtual-machines/windows/>

- Microsoft. (2020c). *Microsoft docs*. Obtenido de Differences between Power Apps and traditional app development approaches: <https://docs.microsoft.com/en-us/powerapps/guidance/planning/app-development-approaches>
- Microsoft. (2020d). *Github*. Obtenido de Descripción de las matrices de puertas programables por campo (FPGA) y procedimiento para realizar la implementación: <https://github.com/MicrosoftDocs/azure-docs.es-es/blob/master/articles/machine-learning/how-to-deploy-fpga-web-service.md>
- Neto, M. V., Amaral, A. D., Silva, N. F., & Soares, A. d. (2020). *Cornell University*. Obtenido de Deep Learning Brasil -- NLP at SemEval-2020 Task 9: Overview of Sentiment Analysis of Code-Mixed Tweets: <https://arxiv.org/abs/2008.01544>
- Newman, D. (2018). *Forbes*. Obtenido de Top 10 Digital Transformation: <https://analyticsconsultores.com.mx/wp-content/uploads/2019/03/Top-10-Digital-Transformation-Trends-For-2019-Daniel-Newman-Forbes-2018.pdf>
- Patwa, P., Aguilar, G., Kar, S., Pandey, S., PYKL, S., Gamback, B., . . . Das, A. (2020). *Cornell University*. Obtenido de SemEval-2020 Task 9: Overview of Sentiment Analysis of Code-Mixed Tweets: <https://arxiv.org/abs/2008.04277>
- Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., & Androutsopoulos, I. (2016). *SemEval-2016 : Semantic Evaluation Exercises*. Obtenido de SemEval-2016 Task 5: <http://alt.qcri.org/semeval2016/task5/>
- Rehkopf, M. (2020). *Atlassian*. Obtenido de Tutorial sobre el diagrama de evolución de Jira: <https://www.atlassian.com/es/agile/tutorials/burndown-charts>
- Repustate. (2020). *Repustate*. Obtenido de Unlock the meaning in your Spanish data.: <https://www.repustate.com/spanish-sentiment-analysis/>
- Sato, K. (2018). *Google Cloud*. Obtenido de What makes TPUs fine-tuned for deep learning?: <https://cloud.google.com/blog/products/ai-machine-learning/what-makes-tpus-fine-tuned-for-deep-learning?hl=es>
- Steen, H., Thurman, P., Farley, P., Kennedy, D., Christiani, T., & Mungi, R. (2020). *Microsoft docs*. Obtenido de How to: Detect sentiment using the Text Analytics API: <https://docs.microsoft.com/en-us/azure/cognitive-services/text-analytics/how-tos/text-analytics-how-to-sentiment-analysis?tabs=version-3-1>

- Stone, Z. (2019). *ProQuest*. Obtenido de Google Cloud Blog - Google: Cloud TPU Pods break AI training records: <https://bv.unir.net:2257/docview/2363328817?pq-origsite=summon>
- Vonage. (2020). *Vonage*. Obtenido de Análisis de sentimiento: <https://www.vonage.co/communications-apis/programmable-solutions/sentiment-analysis/>
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2019). *Cornell University*. Obtenido de XLNet: Generalized Autoregressive Pretraining for Language Understanding: <https://arxiv.org/abs/1906.08237>
- Zhu, X. (2020). *Github*. Obtenido de Tensorflow Pretrained BERT on Azure Machine Learning: <https://github.com/microsoft/AzureML-BERT/blob/master/finetune/TensorFlow/notebooks/Tensorflow-BERT-AzureML.ipynb>

Anexos

Anexo I. Artículo

Automatización de tareas de validación por medio de procesamiento de lenguaje natural

Edgar Emiro Alvarado González
Máster Universitario en Ingeniería de Software y Sistemas Informáticos
Universidad Internacional de la Rioja – UNIR

ee.alvarado.g@gmail.com

Abstract- This work contains the planning of a project to automate the sentiment analysis validation tasks in service communications with users. The natural language tools currently available to solve this type of problem are analyzed, finding in the SemEval Workshop a good reference of models with references to compare the results obtained. Subsequently, the alternatives at the Software and Hardware level for implementation are analyzed, finding that it is convenient to use an approach of using cloud computing services. Subsequently, the planning elements are raised based on the new process suggested for the use of the tool, with the aim of recognizing modules, phases, user stories and activities for planning, and design elements. For planning, the required work team, schedule, budget, risks and lessons learned are identified. Next, the design elements of the solution are added with UML schematic diagrams, solution architecture, user interfaces, techniques, tools, implementation languages, tasks for fine tuning the model to use, tests to be carried out and usability validation. It is concluded that the objectives of model validation, review of tools in cloud service providers, and workflow planning are met. The importance of establishing the organizational process to be implemented based on the new roles that will perform more analytical than operational tasks is also concluded to give a complete perspective of the customer's contacts with the organization, and that planning can be flexible for adaptation. to particular types of business.

Keywords: Natural language processing, SemEval, model, cloud computing, sentiment analysis.

I. INTRODUCCIÓN

La gran mayoría de organizaciones atienden usuarios. Es importante la conexión de los altos ejecutivos con la opinión de los usuarios, para entender como se está desempeñando la cadena de valor de la organización para solucionar las necesidades de los usuarios en procesos como venta y posventa. Las técnicas de procesamiento de lenguaje natural permiten alcanzar objetivos de automatización y análisis en las comunicaciones a usuarios que permiten tener una perspectiva global de las comunicaciones modelando lo que es importante para una organización en particular.

II. CONTEXTO Y ESTADO DEL ARTE

En la academia existen avances considerables en herramientas de inteligencia artificial para enfrentar desafíos cada vez más complejos de procesamiento de lenguaje natural. Con el avance de las capacidades computacionales se ha logrado consolidar modelos de redes neuronales que pueden compararse por medio de tareas estándar para conocer de forma cuantitativa la precisión en el procesamiento de lenguaje natural. Existe un *Workshop* internacional de Evaluación Semántica llamado *SemEval* (Manandhar, et al., 2014b). En este evento anual se plantean tareas de análisis de lenguaje por medio de herramientas de automatización. Uno de los retos que plantea el presente trabajo

es abordar el análisis de sentimiento automatizado en la validación de comunicaciones corporativas.

En la tarea 4 de *SemEval* 2014 se plantea predecir el sentimiento de aspectos con datos de entrenamiento como los siguientes.

```
<?xml version="1.0" encoding="UTF-8" ?>
<sentence id="813">
  <text>All the appetizers and salads were fabulous, the steak was mouth watering and the pasta was delicious!!!</text>
  <aspectTerms>
    <aspectTerm term="appetizers" polarity="positive" from="8" to="18"/>
    <aspectTerm term="salads" polarity="positive" from="23" to="29"/>
    <aspectTerm term="steak" polarity="positive" from="49" to="54"/>
    <aspectTerm term="pasta" polarity="positive" from="82" to="87"/>
  </aspectTerms>
  <aspectCategories>
    <aspectCategory category="food" polarity="positive"/>
  </aspectCategories>
</sentence>
<sentence id="1579">
  <text>And really large portions.</text>
  <aspectTerms>
    <aspectTerm term="portions" polarity="positive" from="17" to="25"/>
  </aspectTerms>
  <aspectCategories>
    <aspectCategory category="food" polarity="positive"/>
  </aspectCategories>
</sentence>
<sentence id="1787">
  <text>Go inside and you won't want to leave.</text>
  <aspectCategories>
    <aspectCategory category="anecdotes/miscellaneous" polarity="positive"/>
  </aspectCategories>
</sentence>
<sentence id="3128">
  <text>Save yourself the time and trouble and skip this one!</text>
  <aspectCategories>
    <aspectCategory category="anecdotes/miscellaneous" polarity="negative"/>
  </aspectCategories>
</sentence>
```

Figura 1. Datos de evaluación de tarea 4 de *SemEval* 2014
(Manandhar, et al., 2014a).

El modelo funciona en idioma inglés, tal como se puede apreciar en la estructura de datos, se detalla en la etiqueta *<text>* la opinión del comensal, y en las etiquetas *<aspectTerm>* y *<aspectCategories>* se detalla cada aspecto con su valoración. Esta información es la entrada del modelo para comparar el resultado con la valoración. Se maneja un mismo formato para los datos de entrenamiento y de evaluación.

Así como en idioma inglés, hay un avance importante en el análisis de sentimiento en otros idiomas, incluso análisis en textos que incluyen varios idiomas. Existe la tarea 9 de *SemEval* 2020 que tiene como objetivo evaluar el sentimiento de tweets con idiomas mezclados, específicamente Hindi e Inglés, o *Hinglish*, y Español e Inglés, o *Spanglish*. Uno de los participantes en esta tarea (Neto, Amaral, Silva, & Soares, 2020) usó la combinación de los resultados d 4 modelos, *MultiFIT* (Sintonía fina del modelo de idioma multilingüe), *BERT* (Representaciones de codificadores bidireccionales de transformadores), *ALBERT* (*BERT* ligero) y *XLNet* (método de preentrenamiento autorregresivo generalizado que permite aprender contextos bidireccionales maximizando la probabilidad esperada sobre todas las permutaciones del orden de factorización (Yang, et al., 2019)). Este modelo alcanzó una calificación F1 de 72.7%.

El modelo *BERT* pre-entrenado puede ser “sintonizado” adicionando una capa en la salida, lo cual puede ser usado en análisis de sentimiento y otras tareas de procesamiento de lenguaje natural. El resultado del ensamble de los modelos es mejor al resultado de cada uno:

Model	F1	P	R	Acc
Ensemble	0.727	0.729	0.726	0.723
XLNet	0.679	0.696	0.692	0.690
ALBERT	0.679	0.684	0.676	0.675
BERT	0.675	0.680	0.672	0.670
MultiFit	0.665	0.665	0.669	0.662

Figura 2. Resultados de uno de los participantes en la tarea 9 de SemEval 2020 (Neto, Amaral, Silva, & Soares, 2020)

La mayor calificación obtenida en esta tarea fue de 80,6% para el data set de Spanglish:

Rank	System	Positive			Neutral			Negative			Avg. F1
		P	R	F1	P	R	F1	P	R	F1	
1	LiangZhao	88.3	92.6	90.4	18.1	20.9	19.4	59.9	39.5	47.6	80.6
2	rachel	89.0	87.7	88.3	16.0	45.1	23.7	65.3	24.5	35.7	77.6
3	asking28	84.5	90.1	87.2	6.1	4.9	5.4	43.5	29.9	35.4	75.6
4	dpalominop	91.6	77.2	83.8	12.7	30.6	17.9	42.9	58.6	49.5	75.5
5	kongjun	87.1	84.6	85.9	11.1	30.1	16.2	56.1	27.2	36.6	75.3
6	HaoYu	92.9	74.0	82.4	11.9	48.1	19.1	55.2	55.0	55.1	75.2
7	Taha	84.7	89.5	87.0	51.9	20.5	29.4	10.4	17.5	13.0	75.1
8	meiyim	93.0	73.3	82.0	12.1	55.8	19.9	57.7	47.1	51.9	74.5
9	Lavinia_Ap	82.0	97.9	89.2	13.8	3.9	6.1	56.0	8.0	14.1	74.4
10	jupitter	93.6	71.8	81.3	11.0	53.9	18.2	58.1	47.9	52.5	73.9
11	tangmen	91.8	72.5	81.0	11.3	55.3	18.8	59.8	41.6	49.0	73.0
12	hermosillo748	85.4	81.3	83.3	7.3	21.8	10.9	54.1	26.6	35.7	72.8
13	harsh_6	87.7	77.9	82.5	9.5	23.3	13.5	36.1	39.1	37.5	72.5
14	francesita	80.9	99.5	89.2	8.7	1.0	1.7	0.0	0.0	0.0	72.2
15	ajason08	90.1	71.0	79.4	8.2	40.3	13.6	54.7	37.5	44.5	71.0
23	Baseline	89.5	63.0	74.0	7.9	49.5	13.6	47.0	31.0	37.4	65.6

Figura 3. Resultados de tarea 9 de SemEval 2020 para el análisis de sentimiento de tweets en “Spanglish” (Patwa, et al., 2020)

Es interesante observar un ejemplo de los datos de entrenamiento para esta tarea, para observar un ejemplo del gran reto que representa la automatización del análisis de sentimiento:

```

meta      27      positive
Manager   lang1
got       lang1
us        lang1
tacos     lang2
and       lang1
everyone  lang1
is        lang1
just      lang1
chilling  lang1
lol       lang1

```

Figura 4. Ejemplo de datos de tarea 9 de SemEval 2020 para el análisis de sentimiento de tweets en “Spanglish” (Patwa, et al., 2020)

En el ejemplo de datos de la figura anterior se observa que la primera columna contiene la frase, la segunda el lenguaje, y al final de la primera fila se da la calificación del sentimiento global de la frase. El objetivo de los modelos es predecir esta calificación.

Para esta tarea, así como para la mencionada anteriormente de análisis de sentimiento basado en aspectos, se encontraron los mejores resultados en modelos basados en *BERT* (Patwa, et al., 2020).

BERT es el modelo de inteligencia artificial creado por Google para entender el lenguaje de las búsquedas de los usuarios (CE Noticias Financieras English, 2019). Google liberó el código fuente para que un usuario pueda crear una variedad de modelos de análisis de texto de nivel del estado del arte (Gupta, 2018).

Para el uso de *BERT* se requiere dos tareas pre-entrenamiento y sintonía fina (Devlin, et al., 2020):

- Pre-entrenamiento: Es un proceso que requiere gran cantidad de procesamiento, pero se hace una vez por cada lenguaje. La gran mayoría de aplicaciones no requieren de una tarea de pre-entrenamiento.

- Sintonía fina: No es un proceso pesado. Una tarea convencional de *NLP* puede tomar una hora en una *Cloud TPU*, y unas pocas horas en una *GPU*. Más adelante se revisará en detalle las opciones de hardware.

Para el uso de *BERT* en idioma español se encuentran las siguientes opciones:

- *BERT* Multilinguaje (Devlin, et al., 2020)
- BETO (Cañete, et al., 2020b)

Teniendo en cuenta lo anterior, en los siguientes capítulos se procede a profundizar en el uso del modelo *BERT* y soluciones complementarias y similares para planificar el presente trabajo.

Una vez explorado el resultado de los modelos en tareas similares, se procede a analizar cómo se implementará teniendo en cuenta experiencias anteriores en la industria y herramientas similares.

En el presente trabajo se explorarán las alternativas sobre los siguientes aspectos de implementación:

- *Hardware*. La investigación se orienta a conocer la característica del Hardware que sea óptimo para la implementación del modelo requerido.

- *Software*. El Software debe permitir implementar los requisitos funcionales y permitir requisitos funcionales a nivel del estado del arte para este tipo de soluciones.

En la industria de *Machine Learning* y *Deep Learning* existen las siguientes alternativas de *Hardware*:



Figura 5. Alternativas de HW para soluciones de ML y DL (Jawandhiya, 2018).

Las *CPUs* son las unidades de procesamiento de nuestros computadores y teléfonos móviles. En estos dispositivos podemos usar múltiples lenguajes de programación. Las *GPUs*, unidades de procesamiento gráfico, son diseñadas para procesar tareas en paralelo, en general, las *GPUs* se pueden desempeñar bien ejecutando algoritmos de *Deep Learning* porque requieren un alto número de cálculos como sumas y multiplicaciones (Jangamreddy, 2019). Los *FPGAs* (matrices de puertas programables por campo) son un tipo de hardware que puede ser programado y reconfigurado usando *HDL* (Lenguaje descriptivo de *Hardware*). Se ha observado que la relación de desempeño sobre energía consumida es más alta que para los *GPUs*, y la flexibilidad del hardware puede tener ventajas a futuro con los modelos de *Machine Learning* (*ML*) y *Deep Learning* (*DL*) que también son cambiantes (Jawandhiya, 2018). Por esta razón varias empresas han usado estos dispositivos como soporte para su estrategia de servicios de *ML* y *DL*, como veremos más adelante. También existen los *ASICs* (circuitos integrados específicos de aplicación). Es la opción menos flexible, pero a la vez tiene el desempeño más alto de las alternativas (Jawandhiya, 2018). Google usa este tipo de implementaciones con su *TPU* (*Tensor Processor Unit*) (Stone, 2019).

En pruebas realizadas con modelos de *DNN* (Redes neuronales artificiales profundas) se identificó que un *TPU* usa menos de la mitad de la potencia de un sistema de *GPU* y tiene mejor latencia, incluso alcanzando tiempos de respuesta 15 veces más rápidos en el procesamiento que un sistema *GPU*, lo cual hace que *TPU* se convierta en un arquetipo para arquitecturas específicas de dominio (Jouppi, et al., 2017). El tiempo de latencia es muy importante para la experiencia del usuario en la aplicación a implementar.

Teniendo en cuenta lo anterior, para usar los modelos de nivel de estado del arte es conveniente usar servicios en la nube.

Entre los modelos soportados por el *TPU* de Google Cloud Platform se encuentra *BERT*. Incluso se encuentra un modelo de “Speech recognition” para analizar la información proveniente de audios.

III. CONJUNTO DE SOLUCIONES SIMILARES O PARCIALES

En general, las propuestas de soluciones similares no incluyen los siguientes aspectos:

- Explicación de modelo usado, o alguna evaluación de *benchmark* tal como *SemEval*.
- Solución de análisis de sentimiento basada en aspectos. Por lo general son soluciones que evalúan el sentimiento de la frase completa.

Soluciones de análisis de sentimiento de audios de *call center* en idioma español

- *Vonage*: Se ofrece análisis de sentimiento en línea por medio de un “bot” para generar alertas. Con los resultados el bot se entrena para mejorar su análisis (Vonage, 2020).
- *Advanced analytics*: Se menciona proceso y análisis de textos y análisis de audios pero no se da mayor detalle del proceso sugerido (Apex, 2020). Esta empresa es partner de *Google Cloud Platform*.

APIs de análisis de sentimiento en español

Las siguientes APIs tienen soporte para análisis de modelo basado en aspectos en idioma español, pero no explican que modelo usan para evaluar la efectividad del modelo sobre un *benchmark* como *SemEval*:

- *Cloudfmrsive* (Cloudfmrsive, 2020). Tiene una ventaja que puede ser usado por un conector de *PowerApps*. La API de análisis de sentimiento analiza la calificación de sentimiento de la frase completa, no basada en aspectos.
- *Meaning Cloud* (MeaningCloud, 2020). La API de análisis de sentimiento de *Meaning Cloud* sí incluye la valoración por aspectos.
- *Microsoft sentiment analysis* (Steen, et al., 2020). Esta API de análisis de sentimiento también está basada en aspectos.
- *Repustate* (Repustate, 2020). Aunque la API de análisis de sentimiento sí evalúa aspectos, es necesario especificar en el consumo cuales son los aspectos requeridos para evaluar.

IV. METODOLOGÍA DE TRABAJO

Se va a hacer uso de técnicas ágiles para el desarrollo de las fases del proyecto. Aunque se espera mantener un fácil entendimiento de la solución y su arquitectura interna, su adaptación a un contexto particular puede necesitar del levantamiento de requisitos personalizados que resulte en zonas de incertidumbre que se van aterrizando con las entregas tempranas de valor de los *sprints*.

V. PROCESO SUGERIDO

Proceso de consolidación de análisis:

El proceso de consolidación tiene los pasos de procesamiento que debe tener una grabación de atención a usuario para pasar a texto, a formato *TF-Record* y análisis de sentimiento basado en aspectos. El usuario de este proceso, o “Analista de consolidación de análisis” debe tener conocimiento en todos los procesos de información para monitorear y dar soporte en línea al proceso completo. También debe tener acceso a la interfaz que permite cambiar los recursos disponibles para el procesamiento de información.

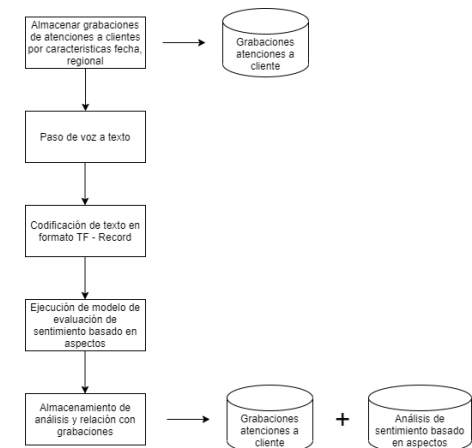


Figura 6. Proceso de consolidación de análisis basado en aspectos (Elaboración propia)

Proceso de validación de información

El “Analista de validación de información” puede ser un analista asignado para revisar el análisis de sentimiento de un servicio y una regional específica, ó un analista ejecutivo que puede tener una perspectiva global y ver todas las regionales, servicios, e información sensible de clientes. Este usuario cuenta con acceso a las perspectivas analíticas y a la reproducción de grabaciones de atenciones específicas.

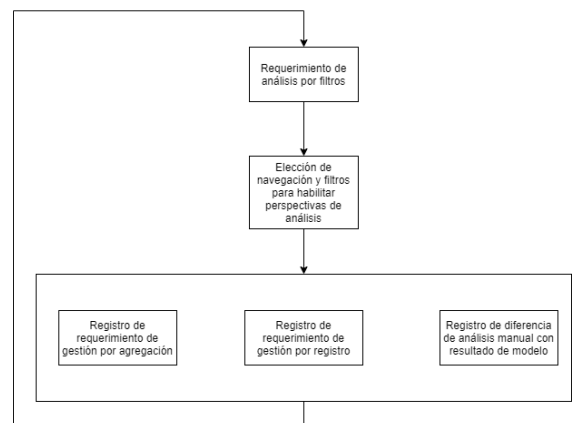


Figura 7. Proceso de analista de validación y gestión de información (Elaboración propia)

Proceso de gestión de información analizada

El proceso es manejado por el rol de “Analista de gestión”. Este rol con base en las variables generales o detalladas identificadas en el proceso de validación es responsable de realizar las siguientes acciones:

- Gestión de procesos de servicio.
- Capacitación a personal de servicio.
- Acciones de personal.
- Gestión con áreas que componen la cadena de valor de la organización.

Proceso de sintonía fina de modelo

El proceso inicial de sintonía fina requiere del levantamiento de requisitos de negocio, y las actividades propias de la sintonía fina del modelo:

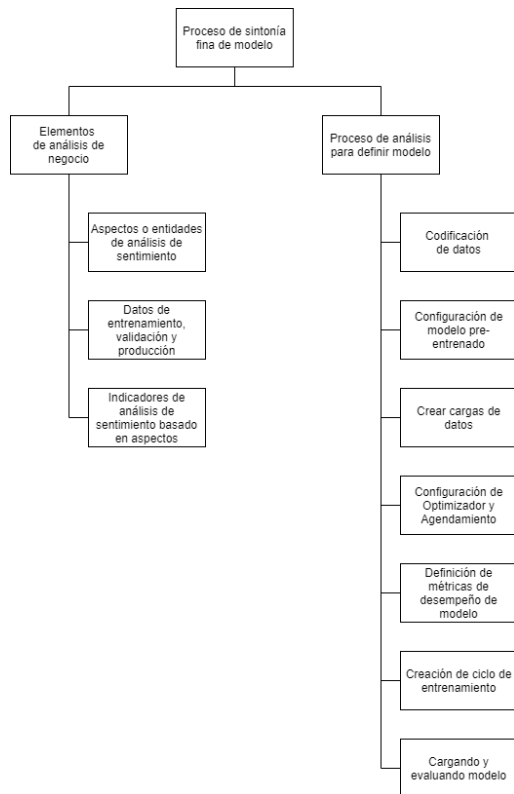


Figura 8. Proceso de sintonía fina de modelo inicial (Google, 2020c)

Una vez el modelo está en producción, se requiere implementar un proceso de sintonía fina frecuente alimentado por la validación manual de analistas. El proceso cuenta con las siguientes fases:

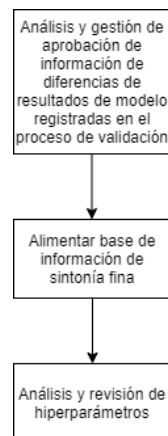


Figura 9. Proceso de sintonía fina frecuente de modelo (Google, 2020c)

Proceso de análisis ejecutivo

Un “Analista ejecutivo” puede realizar las siguientes acciones:

- Análisis de sentimiento basado en aspectos global o segmentado por servicio, regional, o segmento de cliente.
- Análisis proceso de consolidación de información.
- Análisis proceso de validación.
- Análisis proceso de gestión.
- Análisis de proceso de entrenamiento de modelo.

VI. PLANIFICACIÓN

Distribución de esfuerzos

Para el desarrollo de las actividades detalladas en el *Visual Story Mapping* se requiere contar con profesionales con los siguientes roles:

- *Product Owner*
- Ingeniero de datos
- Diseñador de soluciones de *Business Intelligence*
- Desarrollador de *Backend* e integración de servicios
- Desarrollador de *Business Intelligence*
- Desarrollador de *Frontend*
- Analista de pruebas
- *Scrum master*

En total se planifica ejecutar el desarrollo en 5 *sprints*, cada uno de dos semanas, con un probable esfuerzo adicional derivado de la validación de usabilidad.

Por cada rol se consideran los siguientes esfuerzos:

Tabla 1. Esfuerzo por rol.

Roles	Horas
<i>Scrum master</i>	400
Desarrollador de <i>Frontend</i>	272
Desarrollador de <i>Backend</i> e Integración	268
Ingeniero de datos	224
Analista de pruebas	192
<i>Product Owner</i>	177
Desarrollador de <i>Business Intelligence</i>	112
Diseñador solución <i>Business Intelligence</i>	40
Total	1.685

Fuente: Elaboración propia.

Análisis de riesgo

Identificación de aspectos

Los aspectos para la medición de sentimiento deben ser identificados correctamente para cada negocio en particular para que los análisis se relacionen directamente con la estrategia. Por esta razón se plantean grupos focales para el levantamiento de requisitos de aspectos con diferentes roles. En todo el ciclo de proyecto se debe monitorear que los aspectos levantados al inicio del proyecto sean bien identificados y que los datos para sintonía fina sean adecuados para que el modelo maneje los aspectos estratégicos para la organización.

Obtención y limpieza de datos

En general los proyectos de análisis de información tienen sus esfuerzos más representativos en esta fase. El proyecto debe apoyarse con las áreas de tecnología y procesos del cliente para apoyar esta labor.

Precisión del modelo

La precisión del modelo depende del manejo de los anteriores dos riesgos.

Resistencia al cambio

Con la propuesta se propone de manera complementaria cambiar el proceso común de empresas de consolidación de indicadores y validación de información. Es importante apoyar el proyecto con agentes de cambio que entiendan las ventajas del nuevo proceso y mostrar entregas tempranas.

Lecciones aprendidas

Tal como lo plantean las técnicas ágiles, el final de cada *sprint* se debe agendar una reunión de retrospectiva para tratar los temas de mejora continua del trabajo del equipo y el cumplimiento de las metas.

VII. DISEÑO DE SOFTWARE

Requisitos funcionales y no funcionales

RF01-El sistema debe pasar la voz a texto de las grabaciones de atención a clientes identificando los roles de cliente y asesor manteniendo la relación con el id del caso. En caso de error debe almacenar y mostrar las excepciones.

RF02-El sistema debe codificar los textos de las grabaciones de atenciones a clientes en formato *TF - Record*. En caso de error debe almacenar y mostrar las excepciones.

RF03-El sistema debe ejecutar el modelo de análisis de sentimiento basado en aspectos en los textos de grabaciones de atenciones a clientes. En caso de error debe almacenar y mostrar las excepciones.

RF04-El sistema debe almacenar los análisis de sentimiento basado en aspectos relacionados con los casos. En caso de error debe almacenar y mostrar las excepciones.

RF05-El sistema debe permitir analizar la información del análisis de sentimiento basado en aspectos por fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes y asesores.

RF06-El sistema debe permitir registrar requisitos de gestión por agregación, es decir, en un filtro de análisis debe permitir registrar un requisito de gestión que aplique a una combinación de parámetros de entrada fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes y asesores. El requisito de gestión aplica a los registros filtrados, no a un registro en particular. El objetivo es gestionar acciones de mejora que apliquen a un comportamiento masivo, o al menos grupal. El sistema debe entregar un número de *ticket* al asesor que lo registra.

RF07-El sistema debe permitir registrar requisitos de gestión por registro. El sistema debe entregar un número de *ticket* al asesor que lo registra.

RF08-El sistema debe permitir diferencias de análisis manual con resultado de modelo. El sistema debe entregar un número de *ticket* al asesor que lo registra.

RF09-El sistema debe permitir registrar gestiones realizadas sobre los *tickets* identificados en las funcionalidades descritas en los requisitos 6, 7 y 8.

RF10-El sistema debe permitir hacer búsquedas de gestiones por los filtros fecha, hora, regional, servicio, aspectos, rangos de evaluación de sentimiento de clientes y asesores con el objetivo de hacer análisis de comportamientos similares o complementarios.

RF11-El sistema debe permitir visualizar indicadores de los procesos de consolidación de información, validación de información, acciones de gestión y sintonía fina de modelo a los roles que tengan el permiso correspondiente.

RNF01-El sistema debe permitir configurar roles para segregar funciones.

RNF02-El sistema debe permitir la funcionalidad y usabilidad definida en las interfaces de usuario con el objetivo de maximizar la facilidad de uso.

RNF03-A nivel de *performance* el sistema debe procesar la información de paso de voz a texto y análisis de sentimiento basado en aspectos de acuerdo al presupuesto disponible para el uso de *TPUs*. Se espera que las grabaciones mínimo sean procesadas en el mismo día calendario.

RNF04-Con el objetivo de mantener altos estándares de seguridad, en los servicios de *cloud* se requiere que para el acceso a datos se use *SSL* con protocolo *TLS 1.2* para transporte y usar cifrado de datos en reposo.

RNF05-Se requiere alta disponibilidad para los servicios de *AWS* y *Azure* para mantener disponibilidad de 99.995% y un *RTO* de máximo 2 minutos.

RNF06-El tiempo de aprendizaje de uso del sistema debe ser de máximo 2 horas en promedio.

RNF07-La tasa de errores cometidos por el usuario deberá ser menor del 1% de las transacciones totales ejecutadas en el sistema.

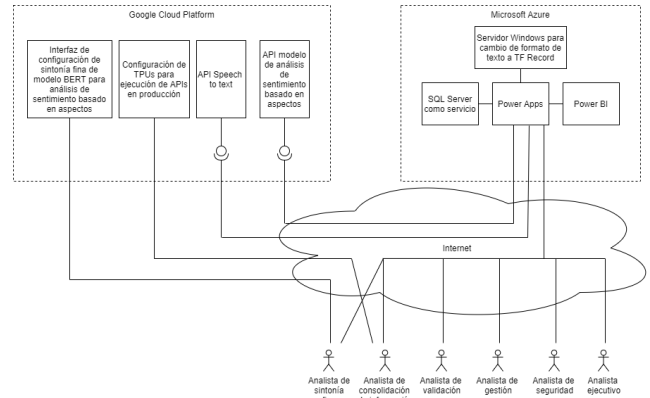
Arquitectura de solución

Figura 10. Diagrama de arquitectura (Elaboración propia)

Interfaz de usuario

Para el proceso de consolidación de procesamiento se tiene las siguientes interfaz de usuario:



Figura 11. Interfaz inicial de proceso de consolidación de información (Elaboración propia)

El sistema permite al usuario navegar del resumen presentado a vistas detalladas por registro para revisar e intervenir el proceso en caso de encontrar errores de procesamiento.

En esta pantalla se muestra el detalle de los registros y el avance en el almacenamiento de la grabación, el paso de voz a texto en formato *TF-Record* y procesamiento de análisis. Se detalla el estado por registro.

Para el proceso de validación de información se tienen las siguientes interfaces:

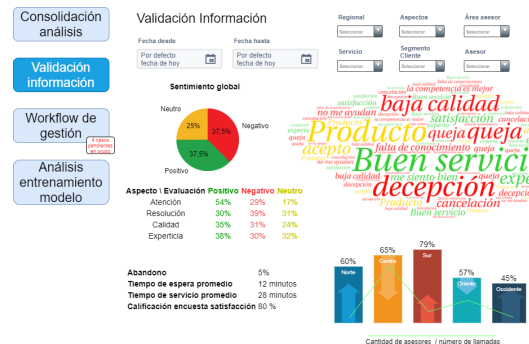


Figura 12. Interfaz inicial de proceso de validación de información (Elaboración propia)

En esta pantalla se muestra una perspectiva general del análisis de sentimiento basado en aspectos para el día actual, el usuario puede cambiar el filtro para ver los datos de un día o rango de fechas diferente. Se permite filtrar por Regional, Servicio, Aspectos, Segmento de clientes, Área de asesor y Asesor. Se muestran indicadores convencionales de servicio, medición de sentimiento global y por aspecto, se muestra por regional el sentimiento global y la tendencia con respecto al día anterior comparado con el indicador de cantidad de asesores por número de contactos. Se muestra una nube de palabras en las que se ven las más usadas por el cliente marcadas con un color que muestra el sentimiento del cliente. Desde esta pantalla se puede seleccionar una regional, un aspecto combinado con la calificación y una palabra para desplegar la siguiente pantalla.



Figura 13. Interfaz que muestra los registros detallados de los casos (Elaboración propia)

En esta pantalla se muestra la información de los registros detallados filtrados en la pantalla anterior. Se puede desplegar u ocultar la información de caso, cliente, asesor y el análisis de aspectos. Al seleccionar un registro, se despliega la siguiente pantalla.

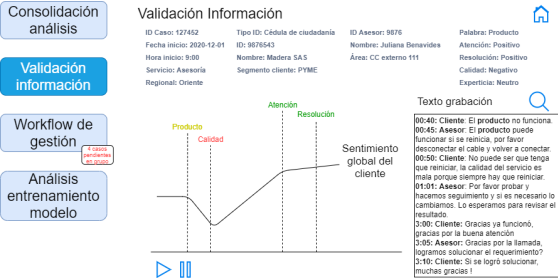


Figura 14. Interfaz de reproducción de grabación (Elaboración propia)

En esta pantalla se permite reproducir la grabación, visualizando el texto. Al reproducir la grabación se va marcando la sección de texto que corresponde. El texto permite buscar palabras. Se muestra las secciones de la llamada en donde se menciona la palabra filtrada, y las calificaciones de aspectos marcadas por colores, dependiendo de la calificación.

Técnicas, lenguajes y herramientas de implementación

El proyecto va a ser implementado usando técnicas ágiles. En el plan se proponen las actividades con base en *sprints* de 2 semanas. En el desarrollo de los *sprints* se manejan las reuniones propuestas por *Scrum*, la reunión diaria, planificación, revisión y retrospectiva. Aunque para el plan se proponen las estimaciones, con el objetivo de tener una perspectiva de tiempo para las entregas, estas estimaciones se realizan con el equipo de desarrollo con la aplicación *Scrum Poker Cards*. La unidad de estimaciones se lleva en horas. El control del tiempo tomado en las tareas y las asignaciones se manejan en *JIRA*. De *JIRA* se consulta el *Burndown Chart* para llevar el control diario del avance y cumplimiento. En la siguiente tabla se muestran los tipos de desarrollo con las herramientas y lenguajes usados.

Tabla 2. Tipos de desarrollos, herramientas y lenguajes de programación.

Tipos de desarrollos del proyecto	Herramientas	Lenguajes
Interfaz de usuario	Power Apps Power BI	.NET T-SQL
Control de roles de seguridad	Power Apps	.NET T-SQL
Workflow de tickets de gestión	Power Apps	.NET T-SQL
Consolidación información de análisis	Azure SQL Server service SSIS - SQL Server Integration Services	T-SQL
Procesamiento y Transformación de información antes de consumo de APIs	.NET Framework	.NET T-SQL
Consumo de APIs Speech to Text y AI Platform	GCP API Speech to Text GCP AI Platform	.NET
Sintonía fina de modelo BERT pre-entrenado	GCP (Google Cloud Platform) Command Line GCP Storage Bucket GCP Cloud TPU GCP Compute Engine VM Github	Python Tensorflow

Fuente: Elaboración propia.

VIII. TAREAS DE SINTONÍA FINA DE MODELO

Las tareas de sintonía fina no hacen parte del sistema de gestión. Son acciones que ejecutan en la implementación y en la sintonía fina sobre el modelo configurado en *Google Cloud Platform*. El sistema interactúa con el modelo para ejecutar los análisis, pero no tiene la funcionalidad de cambiar los parámetros de sintonía fina, estos cambios se deben hacer directamente en la interfaz de *Google Cloud Platform*.

Definición de Requisitos

- Definición de aspectos o entidades de análisis
 - Definición de datos de sintonía fina, validación y producción
 - Definición de indicadores de análisis de sentimiento basado en aspectos
- Análisis y diseño**

- Análisis y diseño de formato y limpieza de datos de sintonía fina, validación y productivos
- Exploración, análisis y perfilamiento de datos y preprocesamiento
- Análisis y diseño de indicadores de análisis de sentimiento basado en aspectos

Implementación

- Codificación de datos
- Configuración de modelo pre-entrenado
- Crear cargas de datos
- Configuración de Optimizador y Agendamiento
- Definición de métricas de desempeño de modelo
- Creación de ciclo de sintonía fina
- Cargando y evaluando modelo
- Implementación de indicadores de análisis de sentimiento basado en aspectos

IX. DISEÑO DE PRUEBAS

El diseño de las pruebas se realizó tomando los elementos de diseño requisitos funcionales y no funcionales, casos de uso, diagrama de componentes. Para cada caso se analizaron los escenarios de las funcionalidades para establecer los roles, resultado esperado y precondiciones de funcionalidades que deben ser ejecutadas antes para obtener el estado en el que la prueba sea válida. Se espera que en la ejecución de pruebas el analista diligencie si el resultado es exitoso y las observaciones para compartirlo con el equipo de desarrollo en caso de requerir ajustes. En total se determinaron 18 casos de prueba funcionales y 4 no funcionales.

X. VALIDACIÓN DE USABILIDAD

Para validar la usabilidad se usa la herramienta *System Usability Scale (SUS)* (Brooke, 2020). En la metodología se propone el uso de una encuesta con resultados en una escala de 0-40 que se multiplican por 2.5 para obtener el resultado de 0-100. El objetivo de la aplicación según el estándar es obtener un puntaje igual o superior a 68.

XI. CONCLUSIONES

En el presente trabajo se ha logrado planificar una solución de validación de información para analizar el sentimiento basado en aspectos de las atenciones a clientes o usuarios, cambiando la definición de las actividades operativas a analíticas. Con la planificación se cumple el objetivo de directores de monitorear todos los contactos de los clientes con la compañía, identificando las opiniones destacadas y habilitando gestiones sobre comportamientos masivos o individuales.

Se validó en fuentes académicas el tipo de modelo a usar para el análisis de sentimiento basado en aspectos como una tarea de procesamiento de lenguaje natural. Se revisan las opciones usadas en soluciones similares de procesamiento de lenguaje natural concluyendo que el modelo *BERT* y sus derivados tienen una mejor precisión para solucionar este tipo de problemas. El rápido avance de este tipo de tecnologías exige una investigación constante para evaluar formas novedosas para abordar este tipo de requisitos.

Se revisaron los proveedores de servicios de computación en la nube, evaluando las mejores opciones a nivel de *hardware* y *software* para la implementación del modelo. Se concluye que actualmente *Google Cloud Platform* ofrece la mejor opción de *hardware* y *software* para usar la *API* de paso de voz a texto y de análisis de sentimiento basado en aspectos.

Como parte de la planificación se incluye el desarrollo de un flujo de trabajo de análisis de información que usa los resultados del modelo para la toma de decisiones estratégicas sobre el proceso de atención a comunicaciones de usuarios, y el planteamiento de un proceso organizacional complementario al funcionamiento de la solución.

En el proceso se plantean roles que al hacer uso de la herramienta hacen parte de la transformación de las actividades y el alcance del análisis, para pasar de análisis de muestras poco precisos, a una perspectiva general de los contactos de todos los clientes para potencialmente analizar todos los canales de comunicación con el cliente, teniendo en cuenta que se analiza voz y texto.

Como parte de la planificación se incluye la toma de requisitos de clientes para la adaptación del modelo a necesidades particulares a nivel de aspectos, terminología, y flujo de trabajo. Se plantea una planificación flexible con uso de técnicas ágiles para personalizar los aspectos a analizar de determinado negocio con potencial de integrar las herramientas convencionales de TI que ya se utilizan para integrar el flujo corporativo.

Se ha concluido que el mayor aporte del presente trabajo es plantear una planificación con un plazo relativamente corto para implementar un proceso organizacional de validación de información y una herramienta que lo apalanca, que cambia la forma en que las empresas perciben las opiniones de sus clientes por canales de atención al cliente. Sería posible abordar objetivos más ambiciosos como:

- Evaluar sentimiento de atenciones ejecutadas por bots.
- Minimizar la interacción humana en procesos repetitivos que puedan ser codificados.
- Complementar la base de información de entrenamiento de modelos de *AI* iterativamente con el conocimiento de expertos abordando los indicadores más relevantes. El uso de herramientas como *Dialogflow* (Google, 2020b) para automatizar la atención a clientes requiere del entrenamiento constante.

- Dar un enfoque a las estrategias de automatización de servicio al cliente orientadas en el sentir del usuario, más que en hacer laberintos de *IVR* con voces automatizadas.

- Automatizar procesos de negocio maximizando la experiencia de los clientes.

Se va a continuar investigando los temas de las tareas del *WorkFlow SemEval* y las herramientas de *cloud computing* para implementar este tipo de soluciones.

REFERENCIAS

- Alfaro, E. (2010). *ConocerElAutor*. Obtenido de El ABC del customer experience: <https://conocerelautor.es/ficheros/doc/00119.pdf>
- Alford, A. (2020). *InfoQ*. Obtenido de Microsoft Open-Sources ONNX Acceleration for BERT AI Model: <https://www.infoq.com/news/2020/01/microsoft-bert-acceleration/>
- Amazon. (2020). *Amazon Web Services*. Obtenido de FPGA-Accelerated Deep-Learning Inference with Binarized Neural Networks: <https://aws.amazon.com/marketplace/pp/Missing-Link-Electronics-Inc-FPGA-Accelerated-Deep/B0784D5WJK>
- Apex. (2020). *Advanced Analytics*. Obtenido de Advanced Speech Analytics: <https://apexinsights.cx/>
- Brooke, J. (2020). *Usability*. Obtenido de System Usability Scale (SUS): <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.htm>
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020a). *dccuchile*. Obtenido de dccuchile / beto: <https://github.com/dccuchile/beto>
- Cañete, J., Chaperon, G., Fuentes, R., Ho, J.-H., Kang, H., & Pérez, J. (2020b). *Google Research*. Obtenido de BETO Examples: https://colab.research.google.com/drive/1uRwg4UmPgYlqGYY4gW_Nsw9782GFJbPt
- CE Noticias Financieras English. (2019). *ProQuest*. Obtenido de How will Google BERT help you find what you're looking for easier?: <https://bv.unir.net:2257/docview/2326418941?pq-origsite=summon>
- Cloudmersive. (2020). *Cloudmersive*. Obtenido de Powerful Natural Language Processing APIs: <https://cloudmersive.com/nlp-api>
- Devlin, J., Markowitz, D., Rao, A., Pu, J., Citro, C., Châtel, G., & Asmro, A. (2020). *Github*. Obtenido de google-research / bert: <https://github.com/google-research/bert/blob/master/README.md>
- Google. (2018). *Google Research*. Obtenido de BERT End to End (Fine-tuning + Predicting) with Cloud TPU: Sentence and Sentence-Pair Classification Tasks: https://colab.research.google.com/github/tensorflow/tpu/blob/master/tools/colab/bert_finetuning_with_cloud_tpus.ipynb#scrollTo=HUBP35oCDmbF
- Google. (2019). *Google Research*. Obtenido de Welcome To Colaboratory: https://colab.research.google.com/notebooks/intro.ipynb#scrollTo=5fCEDCU_qrC0
- Google. (2020a). *Google Cloud*. Obtenido de Google Cloud Pricing Calculator: <https://cloud.google.com/products/calculator>
- Google. (2020b). *Google Cloud*. Obtenido de Dialogflow: https://cloud.google.com/dialogflow?utm_source=google&utm_medium=cpc&utm_campaign=latam-CO-all-es-dr-BKWS-all-all-trial-b-dr-1009133-LUAC0009159&utm_content=text-ad-none-any-DEV_c-CRE_436253090535-ADGP_BKWS%20%7C%20Multi%20~%20Machine%20Learning%20%7C%20D
- Google. (2020c). *Google Cloud*. Obtenido de BERT FineTuning with Cloud TPU: Sentence and Sentence-Pair

- Classification Tasks:
<https://cloud.google.com/tpu/docs/tutorials/bert>
- Gupta, S. (2018). *ProQuest*. Obtenido de Google Open Sources
its BERT Model Source Code:
<https://bv.unir.net:2257/docview/2138312944?pq-origsite=summon>
- Jangamreddy, N. (2019). *Research Gate*. Obtenido de A Survey
on Specialised Hardware for Machine:
https://www.researchgate.net/profile/Nikhil_Reddy7/publication/334647302_A_Survey_on_Specialised_Hardware_for_Machine_Learning/links/5d3842b3299bf1995b457c8d/A-Survey-on-Specialised-Hardware-for-Machine-Learning.pdf
- Jawandhiya, P. (2018). Hardware Design for Machine.
International Journal of Artificial Intelligence and Applications (IJAIA), Vol.9, No.1, 22.
- Jouppi, N. P., Young, C., Patil, N., Patterson, D., Agrawal, G., Bajwa, R., . . . Dau, M. (2017). *Research Gate*. Obtenido de In-Datcenter Performance Analysis of a Tensor Processing Unit:
https://www.researchgate.net/profile/Norman_Jouppi/publication/319856813_In-Datcenter_Performance_Analysis_of_a_Tensor_Processing_Unit/links/599928e60f7e9b3edb16779f/In-Datcenter-Performance-Analysis-of-a-Tensor-Processing-Unit.pdf
- Lappemana, J., Clark, R., Evans, J., Sierra-Rubiac, L., & Gordon, P. (2020). *Science Direct*. Obtenido de Studying social media sentiment using human:
<https://www.sciencedirect.com/science/article/pii/S2215016120300868>
- Li, Q., & Ping, D. (2020). *Amazon Web Services*. Obtenido de Fine-tuning a PyTorch BERT model and deploying it with Amazon Elastic Inference on Amazon SageMaker: <https://aws.amazon.com/blogs/machine-learning/fine-tuning-a-pytorch-bert-model-and-deploying-it-with-amazon-elastic-inference-on-amazon-sagemaker/>
- Li, Q., Li, S., Zhang, S., Hu, J., & Hu, J. (2019). *MDPI*. Obtenido de A Review of Text Corpus-Based Tourism Big Data Mining:
<https://www.mdpi.com/2076-3417/9/16/3300>
- Lixiang, L., Zhigang, G., Zhiming, W., & Liang, Y. (2019). *Alibaba Cloud*. Obtenido de Perseus-BERT: Industry-leading BERT Training Solution Optimized for Performance:
https://www.alibabacloud.com/blog/perseus-bert-industry-leading-bert-training-solution-optimized-for-performance_594717
- Manandhar, S., Androutsopoulos, I., Galanis, D., Papageorgiou, H., Pavlopoulos, J., & Pontiki, M. (2014a). *SemEval-2014*. Obtenido de SemEval-2014 Task 4:
<http://alt.qcri.org/semeval2014/task4/data/uploads/restaurants-trial.xml>
- Manandhar, S., Androutsopoulos, I., Galanis, D., Papageorgiou, H., Pavlopoulos, J., & Pontiki, M. (2014b). *SemEval-2014 Task 4*. Obtenido de SemEval-2014 : Semantic Evaluation Exercises:
<http://alt.qcri.org/semeval2014/task4/>
- Matarranz, A. (2017). *Youtube*. Obtenido de Cuando usar las diferentes herramientas de Analítica de Texto - MeaningCloud:
<https://www.youtube.com/watch?v=6mj96mZJZXE>
- MeaningCloud. (2020). *Meaning Cloud*. Obtenido de Análisis de Sentimiento:
<https://www.meaningcloud.com/es/productos/analisis-de-sentimiento>
- Microsoft. (2020a). *Microsoft Azure*. Obtenido de Azure SQL Database pricing: <https://azure.microsoft.com/en-us/pricing/details/sql-database/single/>
- Microsoft. (2020b). *Microsoft Azure*. Obtenido de Windows Virtual Machines Pricing:
<https://azure.microsoft.com/en-us/pricing/details/virtual-machines/windows/>
- Microsoft. (2020c). *Microsoft docs*. Obtenido de Differences between Power Apps and traditional app development approaches:
<https://docs.microsoft.com/en-us/powerapps/guidance/planning/app-development-approaches>
- Microsoft. (2020d). *Github*. Obtenido de Descripción de las matrices de puertas programables por campo (FPGA) y procedimiento para realizar la implementación:
<https://github.com/MicrosoftDocs/azure-docs.es-es/blob/master/articles/machine-learning/how-to-deploy-fpga-web-service.md>
- Neto, M. V., Amaral, A. D., Silva, N. F., & Soares, A. d. (2020). *Cornell University*. Obtenido de Deep Learning Brasil -- NLP at SemEval-2020 Task 9: Overview of Sentiment Analysis of Code-Mixed Tweets:
<https://arxiv.org/abs/2008.01544>
- Newman, D. (2018). *Forbes*. Obtenido de Top 10 Digital Transformation:
<https://analyticsconsultores.com.mx/wp-content/uploads/2019/03/Top-10-Digital-Transformation-Trends-For-2019-Daniel-Newman-Forbes-2018.pdf>
- Patwa, P., Aguilar, G., Kar, S., Pandey, S., PYKL, S., Gamback, B., . . . Das, A. (2020). *Cornell University*. Obtenido de SemEval-2020 Task 9: Overview of Sentiment Analysis of Code-Mixed Tweets:
<https://arxiv.org/abs/2008.04277>
- Pontiki, M., Galanis, D., Papageorgiou, H., Manandhar, S., & Androutsopoulos, I. (2016). *SemEval-2016 : Semantic Evaluation Exercises*. Obtenido de SemEval-2016 Task 5:
<http://alt.qcri.org/semeval2016/task5/>
- Rehkopf, M. (2020). *Atlassian*. Obtenido de Tutorial sobre el diagrama de evolución de Jira:
<https://www.atlassian.com/es/agile/tutorials/burndown-charts>
- Repustate. (2020). *Repustate*. Obtenido de Unlock the meaning in your Spanish data.:
<https://www.repustate.com/spanish-sentiment-analysis/>
- Sato, K. (2018). *Google Cloud*. Obtenido de What makes TPUs fine-tuned for deep learning?:
<https://cloud.google.com/blog/products/ai-machine-learning/what-makes-tpus-fine-tuned-for-deep-learning?hl=es>
- Steen, H., Thurman, P., Farley, P., Kennedy, D., Christiani, T., & Mungi, R. (2020). *Microsoft docs*. Obtenido de How to: Detect sentiment using the Text Analytics API:
<https://docs.microsoft.com/en-us/azure/cognitive-services/text-analytics/how-to/text-analytics-how-to-sentiment-analysis?tabs=version-3-1>
- Stone, Z. (2019). *ProQuest*. Obtenido de Google Cloud Blog - Google: Cloud TPU Pods break AI training records:
<https://bv.unir.net:2257/docview/2363328817?pq-origsite=summon>
- Vonage. (2020). *Vonage*. Obtenido de Análisis de sentimiento:
<https://www.vonage.co/communications-apis/programmable-solutions/sentiment-analysis/>
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., & Le, Q. V. (2019). *Cornell University*. Obtenido de XLNet: Generalized Autoregressive Pretraining for Language Understanding:
<https://arxiv.org/abs/1906.08237>
- Zhu, X. (2020). *Github*. Obtenido de Tensorflow Pretrained BERT on Azure Machine Learning:
<https://github.com/microsoft/AzureML-BERT/blob/master/finetune/TensorFlow/notebooks/Tensorflow-BERT-AzureML.ipynb>