GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD
Escuela de organización industrial

Máster en Business Intelligence y Big Data, Madrid, 2018/2019

SUMADOCASI

SOLUCIONES BI & BIG DATA “SUMADOCASI"

Proyecto Fin de Máster

Autores:

Abraham Suero Álvarez

Michael Steven Madera Ferreiras

Rafael Miguel Dorville Collado

Roderick Samir Canario Montás

Vigor Situ Lou

Tutor:

Teresa Huerga

Tabla de Contenido

1 - Definición del problema ---- 5
1.1 Introducción ---- 5
1.2 Planteamiento del problema ---- 6
1.2.1 Contexto ---- 6
1.2.2 Objetivo y oportunidad de negocio ---- 8
1.2.3 Motivación y viabilidad de la propuesta ---- 9
2 - Investigación y toma de datos ---- 12
2.1 Validación de Hipótesis ---- 12
2.1.1 Identificación de hipótesis a validar ---- 12
2.1.2 Proceso de validación (entrevistas y encuestas) ---- 13
2.1.3 Análisis y conclusiones de las entrevistas ---- 18
2.2 Análisis preliminar de los datos ---- 20
2.3 Análisis de entorno ---- 21
2.4 Análisis competitivo ---- 22
3 - Análisis y Diagnóstico ---- 24
3.1 Diagnóstico ---- 24
3.2 Estudio de mercado ---- 27
3.3 Análisis de la competencia ---- 28
3.3.1 Competencias Directas ---- 28
3.3.2 Competencias Indirectas ---- 28
3.4 Análisis DAFO ---- 29
4 - Modelo de Negocios ---- 31
4.1 Plan Estratégico ---- 31
4.2 Propuesta de valor ---- 32
4.2.1 Socios clave ---- 32
4.2.2 Recursos clave ---- 33

4.2.3 Relación con clientes ---- 34
4.2.4 Canales de distribución ---- 35
4.2.5 Actividades clave ---- 35
4.2.6 Estructura de coste ---- 36
4.2.7 Fuentes de ingresos ---- 36
4.2.8 Misión, Visión y Objetivos ---- 37
5 - Plan de acción ---- 39
5.1 Alcance del proyecto ---- 39
5.1.1 Objetivos específicos ---- 39
5.1.2 Métricas ---- 40
5.2. Análisis de actividades y tareas ---- 41
5.2.1 Modelo lógico ---- 41
5.2.2 Fase desarrollo de la solución propuesta ---- 41
5.2.3 Arquitectura técnica de la solución propuesta ---- 51
5.3 Análisis de recursos ---- 52
5.3.1 Talento humano ---- 52
5.3.2 Recursos físicos ---- 55
5.4 Gestión del tiempo (cronograma) ---- 55
5.4.1 Listado de tareas ---- 56
5.4.2 Diagrama de Gantt ---- 56
6 - Optimización de los resultados ---- 58
6.1 Proyecto de optimización ---- 58
6.1.1 Beneficios tangibles ---- 58
6.1.2 Beneficios intangibles ---- 60
6.1.3 Beneficios estratégicos ---- 60
6.2 Análisis de rentabilidad del proyecto ---- 61
7 - Conclusiones ---- 64
Anexo A - Resultado de Entrevistas ---- 66
Anexo B - Resultado de Modelos predictivos ---- 86

Capítulo 1 - Definición del problema

1. Definición del Problema

1.1 Introducción

El reto principal de cada negocio es siempre mantenerse a la vanguardia de la tecnología, y más si de quien hablamos es una empresa del sector TELCO (Gran Empresa de Telecomunicaciones). Para lograr lo anterior es necesario formar parte de la era de la transformación digital, tomar decisiones en base a los datos con la finalidad de mantener al cliente con una alta satisfacción en torno a los servicios brindados.

Para estas compañías es de suma importancia que el NPS (*Net Promoter Score*), que es el indicador el cual evalúa la satisfacción de los clientes, siempre esté lo más alto posible. Sin embargo, en los últimos meses estas empresas han tenido múltiples averías importantes en reiteradas ocasiones lo que ha disminuido este indicador, aumentando al mismo tiempo la tasa de cancelación (*churn rate*) y acarreando altas multas y remuneraciones pagaderas al cliente obligatorias equivalente al tiempo sin servicio. Estas medidas son aplicadas a través de INDOTEL, quien funge como órgano regulador para este sector.

Por desgracia, actualmente en la República Dominicana no existe una manera de predecir cuando los equipamientos de un distrito o celdas están próximos a mantenimiento, si necesitan ser reemplazados o si están siguiendo un comportamiento que indica una posible falla que afecte su localidad. Esto, evidentemente se debe a que en la actualidad no se están aprovechando la tonelada de datos que generan los equipamientos a través de los sensores de interconexión que utilizan IOT.

Respondiendo a la necesidad expuesta anteriormente, hemos ideado una solución de Big Data que consiste en el desarrollo de un modelo predictivo de las averías provocadas en los hardware en base a los históricos generados por los sensores que generan datos en tiempo real. Esto estará conectado a una plataforma que permitirá visualizar análisis sobre el estado de los equipos con un fácil acceso y entendimiento para el equipo técnico y aquellos que no manejan un alto nivel de especialización. Como puede ver esta solución, permite tener un mayor control sobre los activos más importantes de la empresa aportándole un gran valor y optimización de sus procesos.

1.2 Planteamiento del problema

WIND TELECOM está presentando averías constantes en la red móvil, específicamente en los elementos que componen las antenas. Dichos elementos son los encargados de brindar el servicio a los clientes. Esta problemática genera inconvenientes internos los cuales provocan una deficiencia al momento de resolver las averías en la empresa.

Estas averías no planificadas generan alertas inesperadas en el departamento de operaciones de la red. Este panorama provoca que dicho departamento se tome un tiempo relativamente considerable en procesar dicha alerta, y obtener los insumos suficientes para poder notificar al departamento correspondiente y poder asignar el recurso adecuado para dar solución a la misma.

Este proceso genera un problema interno con respecto a la mala selección de los recursos debido a que cada recurso posee una agenda de trabajo preestablecida por bloques de tiempo la cual se ve afectada por estas averías, dando lugar a que muchas veces los recursos que son idóneos para resolver una situación no puedan ser utilizados. Esto sucede especialmente cuando un recurso humano se encuentra fuera de la institución dando algún soporte a un cliente, o sino sacrificando otras tareas previamente planificadas para dar solución a la problemática inesperada, generando así costos operativos adicionales a la empresa.

Adicional a esto, la mala gestión de recursos está afectando al departamento de compras y contrataciones de la empresa debido a que este departamento depende completamente de los lineamientos trazados por el departamento de mantenimiento para la asignación de recursos y comprar elementos para tener disponibles para resolver las averías. Estos componentes juntos provocan al final una mala experiencia al cliente, lo cual se pudiera traducir en una deserción de los clientes.

1.2.1 Contexto

Este proyecto se estará realizando en una primera etapa para la empresa de WIND TELECOM, S. A. Esta es una empresa de telecomunicaciones en República Dominicana, cuyo enfoque es brindar servicios integrados de Internet, Televisión en línea y Teléfono a sus clientes (público general),

así como servicios especializados a empresas, instituciones y a otras prestadoras de servicios, garantizando calidad, costo y entrega. Su red de voz y datos es completamente IP; además está basada primordialmente en tecnología LTE LTE, apoyándose también en accesos punto a punto por microondas y redes de fibra óptica para el transporte de grandes volúmenes de información.

Su misión es ofrecer soluciones tecnológicas, innovadoras, ágiles y flexibles enfocadas en la satisfacción del cliente a través de un equipo humano altamente capacitado, con pasión e integridad obteniendo resultados que aportan valor a los diferentes grupos de interés manteniendo un alto sentido de responsabilidad corporativa.

Su visión es ser el proveedor favorito de soluciones de nueva generación para VOZ, INTERNET, CONECTIVIDAD Y ENTRETENIMIENTO en los mercados personales e institucionales satisfaciendo las expectativas de nuestros clientes en Calidad, Costo y Entrega apoyando de manera directa la educación en la RD.

Los valores de la empresa seleccionada son:

- **Flexibilidad:** nos adaptamos con facilidad a nuevos entornos, requerimientos y necesidades de nuestros clientes, empleados y accionistas. Somos una empresa con gran apertura y fluidez para responder con agilidad a los constantes cambios y tendencias del mercado.
- **Innovación**: somos una empresa que promueve la creatividad y proactividad. Buscamos constantemente hacer las cosas de manera diferente para agregar valor a nuestros clientes a través de nuevos productos y servicios.
- **Escuchar la voz del cliente**: conscientes de que el cliente es lo primero, tenemos una actitud de escucha y cuidado de nuestra relación con él para satisfacer y anticipar sus necesidades.
- **Enfoque en resultados**: todas nuestras acciones y metas a corto y largo plazo buscan siempre maximizar el valor de la empresa para el accionista. Buscamos eficiencias de una manera constante en nuestras acciones del día a día, proyectos y procesos.
- **Desarrollo de Empleados**: cuidamos nuestra gente. Reconocemos y apoyamos el desarrollo de nuestro personal a través de programas de entrenamiento continuo, involucramiento y un liderazgo participativo.

- **Pasión**: somos apasionados con lo que hacemos y esto nos lleva a ser ágiles y actuar con sentido de urgencia y responsabilidad, cumpliendo con los compromisos asumidos con nuestros clientes, empleados y accionistas.
- **Buen Ciudadano Corporativo**: estamos comprometidos con nuestra sociedad y contribuimos con el desarrollo integral de nuestro país. Creemos en el liderazgo que se enseña con el ejemplo.
- **Integridad**: como empresa y como personas actuamos con apego a la ley y a la ética empresarial con respeto de las diferencias.

1.2.2 Objetivo y oportunidad de negocio

SUMADOCASI tiene el objetivo general de mejorar los sistemas de servicios que ofrecen las empresas de telecomunicaciones en la República Dominicana, a través de la implementación de herramientas que permitan tomar decisiones sustentadas en los datos. Inicialmente nos estaremos enfocando en las empresas del sector telecomunicaciones. Uno de nuestros principales atractivos es que, para la implantación de nuestra solución utilizaremos los propios recursos tecnológicos disponibles en cada empresa. De tal manera de que las empresas que reciban nuestros servicios sean impactadas experimentando las siguientes mejoras:

- Reducción de Costos.
- Incremento en la Satisfacción de los clientes finales, Eficientizarían de las actividades de mantenimiento.
- Redefinición de los sistemas preventivos de mantenimiento.
- Incremento en la comprensión del negocio.

Estos beneficios planteados se pretenden alcanzar por medio de la implementación de Tecnologías de Big Data de múltiples fuentes, a través de las cuales se recopilan información para que posteriormente las mismas sean consumidas, analizadas y presentadas antes los interesados a través tecnologías de 'Inteligencia de Negocios'.

SUMADOCASI - WIND Telecom. Nuestra empresa ofrece la implementación de una herramienta que le permita solucionar o minimizar las apariciones de las fallas en los equipos de la red LTE, que sirven para brindar los servicios de internet, voz y datos a sus clientes finales. La

solución consiste en la implementación de un sistema predictivo de fallas, con el cual se puedan "erradicar" los constantes mantenimientos correctivos y se pueda calendarizar efectivamente el plan de mantenimiento preventivo en dicha infraestructura.

1.2.3 Motivación y viabilidad de la propuesta

Al pasar de los años, la tecnología ha sido uno de los factores más desarrollados, desde simples tareas informáticas, la robótica y la electrónica, trayendo como consecuencia grandes aportes a la humanidad, mejorando la calidad y tiempo de las actividades y tareas del pasado, que eran trabajadas de manera manual.

Estos aspectos vanguardistas que hoy en día forman parte importante de nuestra vida van sufriendo múltiples mejoras e innovaciones y gracias a ello la sociedad va contando con avances tecnológicos que llevan a descubrimientos de herramientas que impactan positivamente en los procesos y conocimientos del ser humano. Una de las herramientas importantes de hoy en día que ha ayudado de múltiples maneras a las empresas en pronosticar e identificar patrones que afectan negativa o positivamente son los diferentes componentes del machine learning.

En la actualidad, encontramos con más frecuencia artículos y estudios de cómo las máquinas pueden evitar sus averías por medio de datos históricos y actuales ayudando a las empresas con algoritmos de machine learning a predecir esas averías en un futuro, la misma adoptada con el término de mantenimiento predictivo, que ha pasado de ser un simple concepto teórico a una parte importante de los procesos de muchas empresas, ayudándoles a definir estrategias para obtener mejores resultados.

Según un estudio de IBM Analytics, las ventajas del Mantenimiento Predictivo son mucho mayores que las de métodos de mantenimiento tradicionales y reportan un retorno de inversión 10 veces mayor, y utilizando Mantenimiento Predictivo se reducen los costes de mantenimiento entre un 20-25% más que con otro tipo de mantenimiento. También se reducen las averías entre un 70-75% más, y los tiempos de paradas entre un 35-45% más.

En República Dominicana, el sector de las telecomunicaciones ha demostrado ser uno de los más competitivos y el que más rápido se adapta e invierte en nuevas tecnologías. Uno de sus principales activos es la calidad con que sirve a los usuarios, sin importar si es telefonía fija, móvil, internet o televisión de paga.

Sin embargo, al paso del tiempo, las problemáticas aumentan generando mayor coste de producción y mano de obra, provocando que cada vez los márgenes de beneficio sean más pequeños, pero con esta oportunidad de mantenimiento predictivo se puede acortar esta brecha, mejorando así la fiabilidad y el desempeño, reduciendo paradas no programadas y costos.

Con todo esto, se puede asegurar que la implementación de mantenimiento predictivo para la empresa WIND TELECOM propone para el sector de las telecomunicaciones, resolver problemas actuales y futuros de las averías.

Capítulo 2 - Investigación y toma de datos

2 - Investigación y toma de datos

2.1 Validación de Hipótesis

Luego de haber planteado las hipótesis iniciales se ejecutaron los procedimientos de validación pertinentes, con la finalidad de avanzar en el diseño del modelo de negocios propuesto y enfocar los esfuerzos donde realmente se necesiten.

2.1.1 Identificación de hipótesis a validar

Las hipótesis planteadas se clasificaron según el origen de las misma para así facilitar la comprensión de cómo impactaría las áreas involucradas por cada una y poder, de mejor forma, escudriñar con gran profundidad cada supuesto.

Hipótesis de cliente

- El cliente inicial es WIND Telecom
- Los clientes internos dentro de WIND Telecom se dividen en dos grupos:
    - Operaciones de la Red:
        - Operaciones y Mantenimiento de la Red
        - Centro de Operaciones de la Red (NOC)
        - Performance de la Red
    - Ingeniería de la Red;
        - Ingeniería y Optimización de la Red LTE
- Entendemos que los clientes les puede interesar un mecanismo que les facilite el cumplimiento de sus objetivos.
- Los clientes potenciales a futuro son: ALTICE, CLARO, VIVA. A estos clientes no se le ejecutará ningún tipo de entrevistas

Hipótesis del problema

- WIND Telecom posee una Tasa de deserción fuera de sus objetivos. Por lo que entiende que mejorar la calidad de experiencia de sus clientes es un punto clave para reducir la misma.
- Los altos costos operativos invertidos en el mantenimiento reactivo evitan que la empresa pueda expandir sus capacidades y ofrecer un servicio de mejor calidad a sus clientes.
- Visibilidad reactiva del performance y uso de la red, esto hace que los tiempos de solución de los problemas se vean seriamente afectados. Lo cual de igual forma compromete la calidad de experiencia de sus clientes.

Hipótesis producto/solución

- Un sistema de predicción de averías beneficiará económicamente a WIND TELECOM, evitando acciones correctivas y sustituyéndolas por preventivas.
- Se aumentará significativamente la satisfacción de los clientes de la empresa al no ser afectados por las celdas averiadas por largos periodos de tiempo.
- Existe disponibilidad de acceso a los datos del comportamiento de los equipos para poder realizar una buena predicción.

2.1.2 Proceso de validación (entrevistas y encuestas)

Para el proceso de validación de las hipótesis se realizaron entrevistas a los departamentos involucrados en el proceso de solución de las averías en la red móvil, con el objetivo de comprender el rol e importancia de cada área en todo el proceso. Esto permitió determinar si la solución planteada en este proyecto será efectiva para el área de negocio de la empresa WIND TELECOM.

Las entrevistas dividieron por áreas. Las áreas seleccionadas fueron:

- Vicepresidencia de Tecnología y Operaciones de la Red
- Dirección de Operaciones y Mantenimiento de la Red
- Dirección de Ingeniería de la Red
- Dirección de Operaciones de Servicios
- Gerencia de Compras y Logística

Luego se procedió a entrevistar al personal técnico de la empresa para verificar sus conocimientos sobre los procesos internos. Y por último se entrevistaron a las posiciones directivas quienes son los que pautan el futuro de la empresa, con la finalidad de que la propuesta planteada forme parte también de lo que quieren alcanzar en la visión empresarial.

Entrevistas

Se realizaron labores de levantamiento de información en 5 áreas distintas de la empresa WIND TELECOM. Estas áreas fueron seleccionadas luego de proceso de investigación donde se determinó que impactan el servicio de normalización de averías en la red móvil ofertada tanto directa como indirectamente. Las áreas resultantes/seleccionadas fueron: Tecnología y Operaciones de la Red, Ingeniería de la Red, Operaciones de la Red, Compras y Logística y Operaciones de Servicios.

El método seleccionado para el levantamiento de información fue a través de entrevistas a personas individuales, una (1) por cada una de las áreas seleccionadas.

Durante el proceso se lograron validar puntos claves que ayudan a tener más certitud sobre la aprobación de las hipótesis planteadas, y ayudan también a conocer quiénes son los diversos actores muestran interés en una solución como la planteada en este proyecto.

Las respuestas otorgadas por los entrevistados se detallan en el ***Anexo 1 - Resultados de Entrevistas para WIND TELECOM***, *donde se encuentran transcritas.*

Resumen de Entrevistas

Entrevista área de Tecnología y operaciones

En el área de Tecnología y operaciones se entrevistó a Patricio Ferreiras, vicepresidente de tecnología y operaciones en las redes, el cual aportó los siguientes puntos claves:

- Manejan indicadores de tipo de servicio y niveles de servicio que brindan o proponen a los clientes
- Producen y consumen la mayor parte de las informaciones manejada en la empresa, obtenida tanto de procesos internos como externos
- Manejan las informaciones en base de datos, Excel y archivos de texto planos

- Las herramientas para realizar sus funciones son ofimáticas, herramienta de manejo de BI y de manera básica Big Data
- Cada fallo de elementos está bajo responsabilidad total del área de tecnología
- Existen informaciones de monitores de estado de salud de los elementos de la red móvil, datos de disponibilidad de los nodos de la red móvil, datos de satisfacción del cliente con el servicio y tasa de deserción (churn) la cual tienen promedio 3 - 5 años de histórico y se puede tener acceso a ellos a través de él.
- Los datos de satisfacción de clientes y tasa de deserción (churn rate) los produce la dirección de Operaciones de Servicio.
- Existen archivos planos con suficientes datos de los costos de operar, de mantener y los costos de actualizar la red móvil, estos datos se tienen con los fines de entender la rentabilidad de los servicios que se monta en esas redes y la rentabilidad de los nodos cuando se construyen. Existe también hoja de cálculo en Excel que contienen información detallada de cada nodo, especialmente de la red móvil
- Apoya la implementación de predicción de fallas, no solo en las averías de la red móvil, sino en otros ámbitos donde hay mucho menos datos automatizados

Entrevista área de Operaciones y Mantenimiento de la Red

En el área de Operaciones y Mantenimiento de la Red se entrevistaron a Elvy Durán, Director de Operaciones de la Red y a Bin Hon NG, Gerente del Centro de Operaciones de la Red (NOC), los cuales aportaron los siguientes puntos claves:

- En el NOC manejan las siguientes informaciones:
    - Averías de todos los elementos de la Red que sirven para prestar servicios masivos y corporativos
    - Disponibilidad de los equipos de la red
    - Operatividad de los equipos de la red
    - Entrada de las averías de los servicios en la red móvil y voz.
- De estas informaciones, solo las siguientes se producen en el NOC:
    - Entrada de averías red móvil y voz
    - Visibilidad de la red
    - Disponibilidad de la red

  - Recepciones de las alertas o alarmas
- Las siguientes informaciones se consumen desde otro departamento: estadísticas performances se obtienen de otro departamento.
- La data que maneja este departamento una parte se consume de base de datos y la otra de archivos en Excel.
- Los datos de mayor impacto se sacan de la herramienta de ZTE y pueden ser exportados a Excel y/o csv.
- Las informaciones son manejadas a través de reportes, administración de base de datos y Power BI
- Apoya la implementación de predicción de fallas, no solo en las averías de la red móvil, sino también en la red de transporte IP o energía eléctrica
- Comenta que con lograr predecir con antelación un 30% o 40% las averías de la red móvil se pueden llegar a reducir significativamente los reportes de quejas de los clientes y reducir en al menos un 30 % el costo operativo

Entrevista área de Ingeniería de la Red

En el área de Ingeniería de la Red se entrevistó a Alan Molina, Gerente de Ingeniería y Optimización de la Red LTE el cual aportó los siguientes puntos claves:

- Funciones del área son:
  - Diseño e implementación de nuevas localidades o nodos en la Red LTE
  - Diseño de los mapas de cobertura de la Red LTE
  - Soporte de 2do nivel de cara a las averías en la red LTE RAN
  - Soporte general de cara a las averías en el CORE
  - Solución y/o normalización de los problemas de performance que presenten los equipos del RAN y CORE de la Red LTE nivel.
- Las informaciones consumidas por el área son las siguientes:
  - KPIs del performance RAN de la Red LTE
- Las informaciones producidas por el área son las siguientes:
  - Datos de configuración de los equipos del RAN y CORE de la Red LTE
  - Datos de los drive-test ejecutados en el RAN de la Red LTE
  - Mapa de cobertura de la Red LTE

- Utilizan herramientas como:
    - PowerBI, Excel para el procesamiento y análisis de la data
    - NMS ZTE (Netnumen RAN y CORE) para configuración de los equipos de la Red LTE
    - TEMS para la ejecución de los drive-test
    - QGIS para el procesamiento de los datos del drive-test
    - PLANET para el diseño de los mapas de cobertura
- Las informaciones producidas y consumidas por el área se almacenan en:
    - Base datos NMS ZTE:
        - Netnumen RAN, Netnumen CORE:
            - Configuración y performance de los equipos
    - Archivos:
        - Drive-test, mapas de cobertura en archivos con formatos de las aplicaciones listadas arriba
        - KPIs del performance RAN de la Red LTE en formato Excel
    - El Netnumen CORE de ZTE solo almacena hasta 3 meses máximo de la data cruda, hay que validar este dato
- Soluciones a principales problemáticas del área:
    - Procesamiento automático de datos de performance para generación de KPIs
    - Proceso para procesar datos de drive-test de manera eficiente
    - Implementación de herramientas o cuadro de mandos para tener visualización, en tiempo real o cercano, de la salud de la red LTE, esto le permitirá mejorar sus tiempos de respuestas ante problemas de performance en los elementos de la red LTE

Entrevista área de Compras y Logísticas

En el área de Compras y Logística se entrevistó a Luisa Pérez, Gerente Senior de Compras y Logística la cual aportó los siguientes puntos claves:

- Funciones del área son:
    - Logística para adquisición de recursos

- Las informaciones consumidas por el área son las siguientes:
    - Órdenes de compra creadas por las áreas solicitantes
- Las informaciones producidas por el área son las siguientes:
    - Órdenes de compra recurrentes para abastecer el almacén
    - Registro histórico de órdenes de compra
- Utilizan herramientas como:
    - Ofimática (Excel y Word)
    - SAP para el manejo y proceso de aprobación de las órdenes de compra
- Las informaciones producidas y consumidas por el área se almacenan en:
    - En SAP
    - Archivos de Excel y Word
- Soluciones a principales problemáticas del área:
    - Incremento licenciamiento y capacidades del SAP
- Los procesos de este departamento representan un 3-4% del costo operativo total de la empresa.
- La adquisición de equipo casi siempre se hace en momentos de crisis, poseen baja disponibilidad de equipos en almacén, esta práctica hace que la empresa tenga que pagar un alto precio por dichos equipos ya que tienen que utilizar flete aéreo lo cual representa el 30% del costo total del equipo

2.1.3 Análisis y conclusiones de las entrevistas

Luego de haber procesado todos los datos arrojados en las entrevistas a las diferentes áreas seleccionadas las cuales se encuentran involucradas en el proceso de normalización de las averías en los elementos de la red móvil vemos como las principales problemáticas de cada área se reducen a cuatro factores fundamentales, ordenados de mayor a menor importancia:

- Reducción en costos operacionales
- Mejora en la satisfacción del cliente con el servicio
- Implementación sistemas autónomos para procesamiento de datos
- Implementación cuadro de mandos para monitoreo eficiente del performance de la red

Con la reducción en costos operacionales la empresa dedicaría estos recursos para la expansión de su red lo que se traduciría en más ingresos y clientes más felices.

Una mejora en la satisfacción del cliente se puede traducir en mayores ingresos, puesto que un cliente feliz atrae a nuevos clientes por medio de referencias.

Con la implementación de sistemas para procesamiento de datos y cuadro de mandos para monitorear el performance de la red móvil se mejorarán tanto la calidad de experiencia del cliente y como los KPIs operacionales de cada área los cuales se traducen a un mejor aprovechamiento de las herramientas disponibles y del presupuesto disponible.

Luego de ver todos los datos manejados en cada área se concluyó que los datos a los que se tendrá acceso serán los siguientes:

- *Churn Rate* detallado Red LTE LTE (Tasa de deserción)
- Encuestas Satisfacción Clientes Red LTE LTE
- Entrada Averías celdas Red LTE LTE
- Disponibilidad celdas Red LTE LTE
- Performance RAN Red LTE LTE
- *Alarm History* Red LTE LTE
- Costos operacionales para solución de averías Red LTE LTE

Es importante destacar que las informaciones facilitadas por WIND TELECOM estarán bajo un acuerdo de confidencialidad que firmarán todos los integrantes de este grupo con la empresa.

En conclusión, con el listado de informaciones mencionado anteriormente se cree que se cuenta con datos suficientes para entrenar el modelo de predicción de averías propuesto y por ende se considera que el proyecto planteado es viable.

2.2 Análisis preliminar de los datos

Para esta fase vamos a respondernos las siguientes interrogantes.

¿De qué datos disponemos?

- Churn Rate detallado Red LTE LTE (Tasa de deserción)
- Encuestas Satisfacción Clientes Red LTE LTE
- Entrada Averías celdas Red LTE LTE
- Disponibilidad celdas Red LTE LTE
- Performance RAN Red LTE LTE
- Alarm History Red LTE LTE
- Costos operacionales para solución de averías Red LTE LTE

¿Qué datos a los que tenemos acceso no se están recogiendo?

- Alarm History Red LTE LTE

  Estos datos solo se almacenan en la plataforma encargada de gestionar y monitorear los elementos del hardware, no se está almacenando histórico de los mismos, solo te tienen datos del 2018 y 2019.
- Costos operacionales para solución de averías Red LTE LTE

  Estos datos solo se almacenan en archivos de Excel, no existe un sistema que tenga todo el histórico unificado.
- Encuestas Satisfacción Clientes Red LTE LTE

  Para estos datos debemos validar la metodología de procesamiento y almacenamiento.

¿Qué datos pueden generarse a partir de nuestros productos y operaciones?

- Predicción averías por elementos
- Set de variables que impactan directamente en el desempeño de su red, así como en la satisfacción de sus clientes

¿Qué datos podríamos obtener de otros que nos serían de utilidad?

- No hemos identificado otros datos hasta ahora.

¿Qué datos que tienen otros podríamos usar en una iniciativa conjunta?

- No hemos identificado otros datos hasta ahora.

¿Cómo podríamos estructurar y analizar nuestros datos para generar mayor valor?

- Creando datamarts departamentales, le servirá para poder efectuar análisis específicos para el área
- Creando dashboard departamentales, le servirá para poder tener una vista general de cómo andan los KPIs departamentales
- Crear dashboard ejecutivos, les servirá para ver cómo anda el negocio, ver dónde pueden mejorar o eficientizar el uso de recursos

¿Estos datos son valiosos internamente para nosotros, o para nuestros clientes actuales, o para nuevos clientes potenciales, o para otras industrias?

Entendemos que son valiosos para nosotros, nuestros clientes actuales y clientes potenciales ya que al manejar datos similares en la industria de las telecomunicaciones podremos sugerirles a nuestros clientes potenciales qué datos podrían generar insight sobre la operatividad, desempeño y calidad de experiencia de los clientes de su empresa.

2.3 Análisis de entorno

Hoy en día se observa una tendencia creciente en el uso de algoritmos de machine learning para la solución a los diversos problemas que presentan las empresas prestadoras de servicios o en la industria de manufactura de cara al análisis de datos para obtener insights sobre el comportamiento de los usuarios o infraestructuras.

Dentro de la industria de telecomunicaciones se está comenzando a implementar tecnologías de machine learning para crear modelos predictivos de deserción de clientes, el cual les permite crear planes de acción preventivos para retener a clientes con una probabilidad de deserción alta, así como también observar comportamientos y patrones en el uso del tráfico hacia internet, esto les permite diseñar planes y servicios para segmentos de clientes específicos que se adapten a sus necesidades.

Observando estas tendencias en el uso de tecnologías de machine learning se ha tomado la decisión de utilizar las mismas para implementar un algoritmo predictivo de averías en el hardware de la red móvil de cualquier empresa de telecomunicaciones. Inicialmente se ofrece esta solución a

WIND telecom con miras de que la misma se reproducida, por el mismo grupo que plantea esta propuesta, en otras empresas de telecomunicaciones en República Dominicana.

2.4 Análisis competitivo

Actualmente no se cuenta con información que valide la existencia de alguna empresa local que brinde este tipo de soluciones. Aunque sí se conoce que WIND Telecom e INTEC tiene acuerdos bilaterales, para el 2017 fruto de esos acuerdos fue implementado un algoritmo predictivo del churn (deserción de clientes) desarrollado por un equipo de estudiantes liderado por un catedrático. Actualmente el mismo equipo que implementó el algoritmo predictivo del churn está desarrollando un algoritmo usando técnicas de Sentiment Analysis para estudiar el comportamiento de sus clientes.

A modo general este grupo no conoce a ninguna empresa local o persona que haya ejecutado un algoritmo para predecir averías del hardware, ni en redes móviles como en ningún otro tipo de sistemas, por lo que entendemos podríamos estar siendo los pioneros en esa área específica.

Capítulo 3 - Análisis y Diagnóstico

3 - Análisis y Diagnóstico

3.1 Diagnóstico

Tal como se puede apreciar en el artículo de Severino (2017), uno de los sectores de mayor crecimiento en las República Dominicana es el sector de las telecomunicaciones pues es uno de los más competitivos y el que más rápido se adapta e invierte en nuevas tecnologías. Viendo que sus principales activos es la calidad con que sirve a los usuarios, a través de todos sus servicios como son la telefonía fija, móvil, internet o televisión de paga.

Según datos del Instituto Dominicano de las Telecomunicaciones (INDOTEL), se puede ver que a la fecha existen unas 76 empresas registradas que ofrecen servicios de internet, televisión por suscripción, telefonía fija y móvil. Sin embargo, actualmente este mercado se encuentra liderado solo por 5 empresas las cuales son conocidas como Claro Dominicana, Altice Dominicana, Tricom, Viva Dominicana y WIND Telecom. De igual forma revisando la memoria institucional de INDOTEL correspondiente al 2018 se puede observar el siguiente panorama:

La memoria indica que "*El servicio de telefonía (voz) está comprendido por las tecnologías de acceso fijo tradicional, fijo IP (Internet Protocol) y móvil. A septiembre, el segmento de las Telefonías presentó una ligera ampliación, principalmente determinada por la evolución de telefonía móvil. Para septiembre del año 2018, el número de líneas activas fue de 10,154,300 en todo territorio nacional, mostrando un leve incremento de 0.6 %, comparado con el cierre del año 2017*".

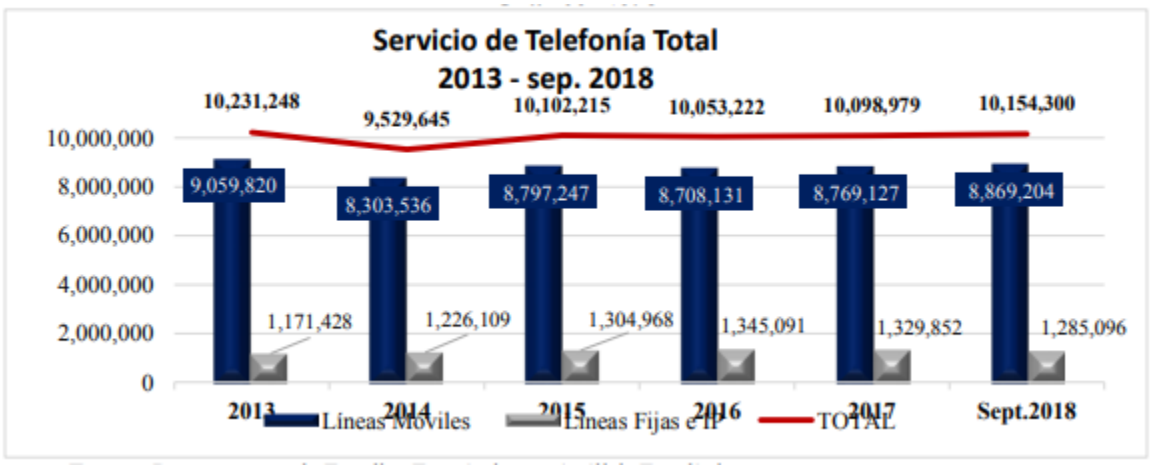


Imagen 1. *Variación Anual de Población de servicios de telefonía en sus distintas categorías, medidos anualmente.*

En donde las líneas móviles poseen un 87.3% del mercado de telefonía del país, con un total de 8,869,204 líneas; el 9.26% corresponde a la telefonía fija tradicional, con 940,711 líneas y el 3.39% restante corresponde a las líneas IP, con un total de 344,385 líneas.

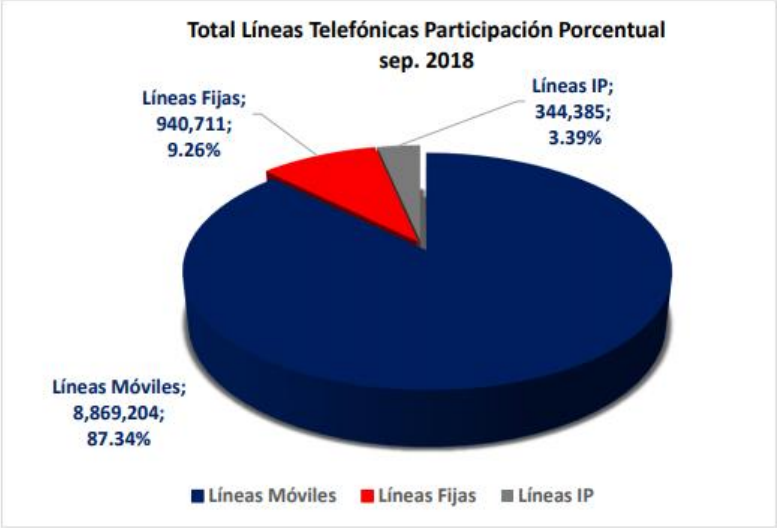


Imagen 2. *Participación Porcentual de los contratos establecidos por tipo de servicios de telefonía.*

Cabe destacar que la República Dominicana cuenta con una Tele densidad de 98.9%, esto quiere decir, que por cada 100 habitantes existen 99 líneas telefónicas en servicio

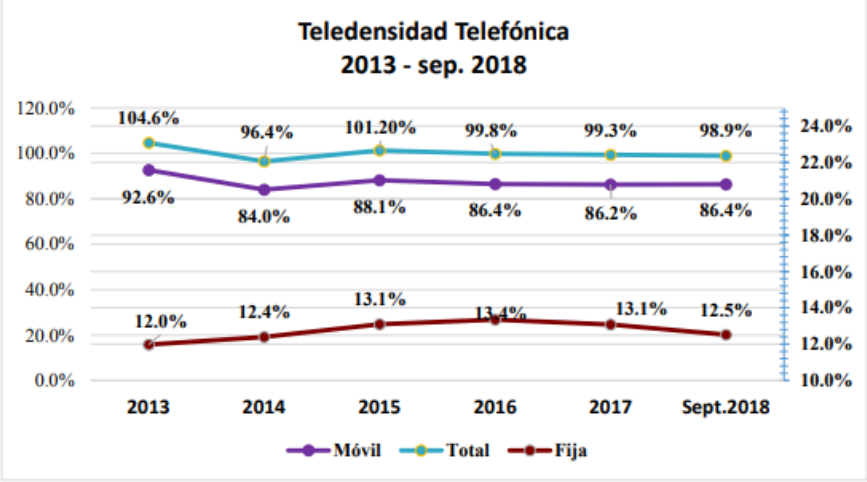


Imagen 3. *Variación porcentual anual para cada tipo de servicio de telefonía.*

En la siguiente grafica podemos apreciar la variación de crecimiento de los servicios telefónicos en los últimos años.

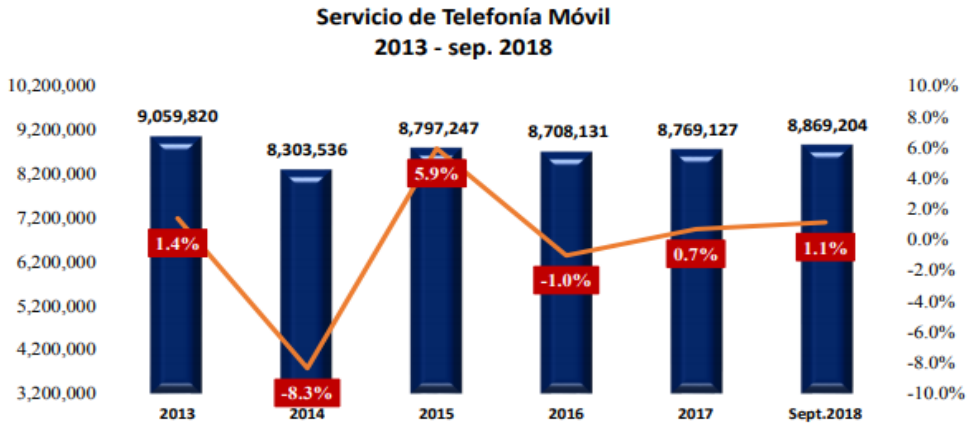


Imagen 4. *Variación porcentual anual Global para los servicios de telefonía.*

Fuente https://indotel.gob.do/media/143560/memoria-institucional-indotel-2018-final.pdf

3.2 Estudio de mercado

El proyecto presta especial atención a los servicios de internet. Importante notar que en la memoria publicada por INDOTEL se valida como el internet constituye una herramienta cada vez más omnipresente, estos indican que "*es una herramienta abierta y rápida que ha transformado el modo en que muchas personas viven, se comunican y hacen negocios, brindando inmensos beneficios a las personas, los gobiernos, las organizaciones y el sector privado. Los resultados revelan que 71 de cada 100 personas tienen acceso a este servicio*".

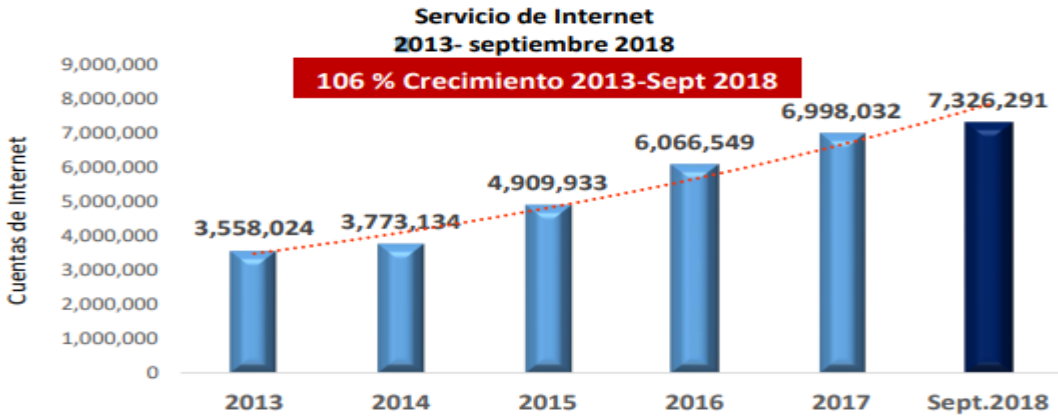


Imagen 5. *Crecimiento anual de cuentas activas que contratan servicios de internet.*

La composición de Mercado Nacional Dominicano de cuentas de internet a septiembre de 2018 estuvo siendo liderado por Claro quien mantiene la primacía al acoger el 55.9% del mercado de República Dominicana, seguido de cerca por Altice, con una participación de 38.8% del total de cuentas, Trilogy con 4.2% y otras prestadoras concretan el 1.1% restante del mercado. Siendo el 0.7% el nicho con el que se estará trabajando en este proyecto.

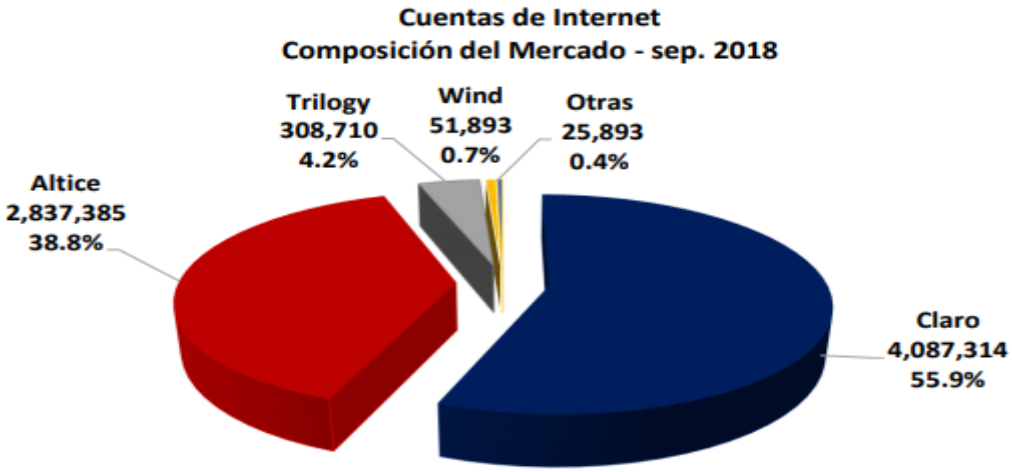


Imagen 6. *Composición del Mercado Nacional de cuentas que contratan Servicios por Internet*

En el mercado del servicio de internet, no se posee una solución de BI para resolver las problemáticas que se presentan, por lo que la solución planteada en este informe sería la primera implementación efectiva en este mercado nacional de la República Dominicana.

3.3 Análisis de la competencia

3.3.1 Competencias Directas

tomando el mercado nacional como referencia para realizar el análisis de competencia, no se encontró ninguna empresa que desarrolle sistemas los cuales tengan como finalidad la predicción de fallos o averías de las empresas de Telecomunicaciones. Este hecho posiciona al equipo desarrollador de este proyecto como la única empresa o grupo en el país especializados en el área mencionada.

3.3.2 Competencias Indirectas

La competencia indirecta sería Iqtek solutions, debido a que poseen servicios de implementación de managed services. Esto es un software de gestión inteligente que permite hacer mantenimientos preventivos tales como: actualizaciones de firmware o notificaciones para hacer actualizaciones en los equipos como limpiezas, cambio de baterías, etc., monitorear todo el sistema y gestionar las alerta. Cabe destacar que este software no tiene habilidades predictivas que gestione las fallas antes de que estas ocurran.

3.4 Análisis DAFO

Nuestro análisis DAFO se centra en los factores internos para reconocer fortalezas y debilidades, y por otra parte nos señalará las oportunidades y amenazas a partir de los factores externos. Se trata de un análisis idóneo y necesario para obtener información crítica a partir de la cual sabremos si actuar o abstenerse de hacerlo.

Nos será más sencillo entender la importancia relativa de algunos de los elementos que en un principio podría parecernos más o menos graves de lo que son. A partir de las oportunidades que tenga nuestro proyecto se establecerá un plan de actuación que deberá pautarse para su cumplimiento, sin perjuicio de futuras revisiones periódicas, por otro lado, necesarias si queremos aumentar nuestras posibilidades de éxito.

***Imagen 7**. Matriz análisis DAFO*

Capítulo 4 –Modelo de Negocios

4 –Modelo de Negocios

SUMADOCASI nace como proyecto final del "Master en Business Intelligence y Big Data" y pone el foco en el sector telecomunicaciones de la republica dominicana.

Este proyecto será el escenario donde afianzar los conocimientos adquiridos y, no menos importante, nos ayudará a reforzar la visión que la innovación basada en datos otorga a las organizaciones un factor estratégico y diferenciador de vital importancia en los tiempos actuales.

El nombre SUMADOCASI proviene de las dos primeras letras de los integrantes del grupo, este nombre nació de forma espontánea en una de nuestras reuniones para dar lluvias de ideas para el tema de proyecto.

SUMADOCASI

Imagen 2. Logotipo de la empresa

4.1 Plan Estratégico

SOCIOS CLAVES	ACTIVIDADES CLAVES	PROPUESTA DE VALOR	RELACION CLIENTES	SEGMENTO CLIENTES
• Gerencia de Tecnología u operaciones • Departamento de Compras y Logística • Departamento de Ingeniería de la Red • Departamento de Operaciones y Mantenimiento de la Red • Proveedores soluciones IoT.	• Gestión de plataforma para manejo confiable de información en tiempo real • Realizar levantamientos de información • Conocer el negocio del cliente • Desarrollo e Implementación de la solución **RECURSOS CLAVES** • Personal capacitado • Sistemas de base de datos • Infraestructura tecnológica para desarrollo solución • Sistemas colección de datos	• Generar conocimiento beneficios para los clientes, sustentados en las buenas practicas • Desarrollar soluciones adaptas a las necesidades del cliente • Reportes en tiempo real y a demanda • Acceso a reportes desde cualquier dispositivo a través de aplicaciones web o móviles • Seguimiento continuo a soluciones implantadas a los cliente	• Relación adaptada al tipo de cliente • Presencial y virtual • Seguimiento y oferta de nuevos productos **CANALES** • Canal propio y directo • Equipo comercial • Redes Sociales, Web y Mailing	• Actual: WIND Telecom, como cliente piloto con el proyecto actual • A futuro: Empresas que deseen mejorar la eficiencia de sus labores diarias a través de una mayor comprensión de su negocio

ESTRUCTURA DE COSTOS	FUENTE DE INGRESOS
• Marketing • Recursos Tecnológicos • Recursos Humanos	• Venta de los servicios ofrecidos • Mantenimiento de la solución tecnológica implementada

Tabla 1. *Business Model Canvas para la Propuesta de Proyecto de SUMADOCASI.*

4.2 Propuesta de valor

La propuesta de valor ofrecida por SUMADOCASI consiste en la creación de conocimiento beneficioso para el cliente, con la finalidad de adaptar las mejores prácticas en el presente basado en un histórico de comportamientos que contempla tanto las herramientas, equipos, y acciones realizadas. A los fines de hacer esto realidad se utiliza una serie de herramientas necesarias (según el tipo de empresa cliente) la cual recopila y analiza de forma automática la historia de eventos sucedidos en un determinado sistema, con la finalidad de presentar el conocimiento de forma utilizable y así facilitar la toma de decisiones inteligentes.

Para este proyecto en particular se le ofrece a WIND Telecom (WT), empresa cliente, la capacidad de mejorar la disponibilidad de los servicios de red móvil que esta ofrece, a través de la creación de un sistema con la capacidad de predecir los mantenimientos necesarios para los hardware. Es importante destacar que dicha oferta impacta de forma directa y positiva a la satisfacción del consumidor del servicio de red móvil, así como también reduce los costos de mantenimientos correctivos.

4.2.1 Socios clave

Al tratarse de un proyecto desarrollado y personalizado, en una primera etapa, para WIND Telecom, éste es el principal partner y patrocinador del proyecto. Durante las fases iniciales se han realizado diferentes acciones con las áreas implicadas:

- **Gerencia de Tecnología y Operaciones de WIND**, los datos que se utilizaran para la creación del modelo predictivo dependen en su totalidad del departamento de Tecnología pues estos serán los encargados de brindar apoyo en el desenvolvimiento y ejecución del proyecto.
- **Departamento de Compras y Logística**, en donde el proyecto tendrá un alto impacto ya que les permitirá hacer gestiones de logísticas más apropiadas para mantenimientos de equipos y a la vez por medio del modelo predictivo podrán tener datos de primera mano de cual equipos habría necesidad de cambiar permitiendo cambiar lo que sucede en la actualidad en donde se enteran de que lo deben cambiar en el momento que sufre un daño mayor.

- **Departamento de Ingeniería de la Red,** la importancia de la solución que le podemos ofrecer a la empresa desde el punto de vista de este departamento radica en la parte de que deben mejorar los KPI, la información de Ingeniería que manejen y para ello propone una implementación de una BD y alguna herramienta para analizar la data para facilitar la información. En la parte de optimización necesitan una herramienta para KPI, pero no tienen los recursos y tiempos para dar respuesta en el tiempo requerido.
- **Departamento de Operaciones y Mantenimiento de la Red,** pudiendo ver que la predicción de fallas tendría un alto impacto en este departamento, pues no solo permitiría manejar mejor las averías de la red móvil, sino también en la red de transporte IP o energía eléctrica. Predecir con antelación un 30% o 40% las averías de la red móvil se pueden llegar a reducir significativamente los reportes de quejas de los clientes y reducir en al menos un 30 % el costo operativo generado en mantenimientos a los equipos.

De igual forma, necesitaremos contar con recursos de apoyo, tales como personas con conocimientos de Inteligencia de Negocios y Big Data que puedan dar soporte al proyecto, por lo que necesitaremos:

- **Proveedores de soluciones IoT:** Empresas que proveen una plataforma como servicio diseñada para proyectos de Internet de las Cosas (IoT) y de máquina a máquina (M2M).
- **Proveedores de desarrollo de aplicaciones**: Empresas que poseen la capacidad para abordar la creación de apps de cualquier complejidad y envergadura, adaptándose a las necesidades de los clientes.
- **Proveedores de IaaS (Infraestructura Como Servicio):** adicional a los recursos de hardware que nos ha proporcionado WIND Telecom, hemos considerado los servicios en la nube de Amazon, debido a la oferta que hace este proveedor en el campo de la analítica de datos y el procesamiento en paralelo.

4.2.2 Recursos clave

Debido a los dos grupos de clientes que queremos atacar debemos poseer una gama recursos que son claves para el buen desarrollo de nuestra propuesta de valor y poder cumplir las necesidades de nuestros clientes

Los recursos claves destacados:

- **Data scientist**, para poder realizar la recopilación, transformación de los datos de manera oportuna y así tener disponibilidad de los mismos
- **Data analyst**, para poder tener una interpretación adecuada de los datos, creando dashboards y diseñando indicadores que permitan una ejecución efectiva de soluciones a los problemas
- **Sistema de base de datos**, para almacenar los datos
- **Sistemas para desarrollo modelo predictivo.**
- **Plataformas de presentación, manipulación y extracción de la información.**
- **Team accionador**, para ejecutar de manera efectiva las acciones sobre las predicciones arrojadas
- **Datos históricos de averías en la red móvil, elementos del RAN.**
- **Datos vida útil de elementos del RAN.**

4.2.3 Relación con clientes

En modo general se mantiene una relación ajustada/personalizada para cada uno de los clientes. Las soluciones ofertadas para los clientes será una adecuada al segmento de empresa al que pertenece y a los servicios o productos que ofrece, con la finalidad de mejorar cualesquiera que sean los beneficios percibidos en la actualidad.

La relación con los clientes en antes y durante de la fase de implementación de la solución ofertada será de forma presencial (face to face), con la finalidad de crear mayores conexiones con el cliente, poder determinar las necesidades reales, y la posición del personal ante el posible cambio que se estará llevando a cabo.

Luego de la implementación de la solución el enfoque entonces cambia. Se torna un 70% de las interacciones serán de forma remota, las cuales se darán con fines de actualizaciones de sistema, ajustes de reportes y revisiones de los algoritmos utilizados para el sistema predictivo.

4.2.4 Canales de distribución

El servicio será comercializado por varios tipos de canales:

- Primero nuestras ventas serán directas al cliente que nos está solicitando la solución por medio de un equipo comercial que se encargará de visitar y resolver la situación planteada.
- Nuestros canales de comunicación estarán fundamentados en la tecnología de la información y comunicación, utilizando:
    - Redes sociales propias y de influencers en los temas de tecnología especialmente en BI.
    - Página web propia de la empresa
    - Comunicación vía correo electrónico (mailing).

4.2.5 Actividades clave

Las actividades claves son la llave para el éxito de este proyecto. Pues a través de estas es que toda la teoría y el conocimiento obtenido se llevará a la práctica. Dentro del grupo de actividades claves identificadas para SUMADOCASI están:

- Mantenerse Actualizado sobre los avances de la tecnología, Big Data & Business Intelligence. Esta actividad es de vital importancia, puesto que desde aquí es que realmente nacen las posibilidades de implementar soluciones altamente eficientes con tecnología de punta, aportando ventaja competitiva al mercado actual.
- Conocer el negocio del cliente. Es importante comprender cuál es el modelo de negocio del cliente, pues solo así se le puede ofrecer un servicio realmente personalizado que le agregue valor al cliente.
- Identificar cuáles son las áreas involucradas en la solución. a pesar de que la solución viene enfocada a un área en específico, la solución propuesta posiblemente necesite consumir o proveer información a otros departamentos o áreas, por lo que es necesario crear el listado de interesados (stakeholders).
- Conocer el sistema de almacenamiento de datos que posee el cliente. Comprender cómo los datos son archivados es crucial para el consumo de los mismos.
- Realizar levantamientos de información de campo. En la mayoría de los casos las empresas no son los que los altos ejecutivos dicen que son, por ende, para poder ofrecer una solución

que se ajuste de forma óptima en sinergia con la realidad operativa y la visión de la empresa, es necesario realizar levantamiento de informaciones a través de entrevistas, encuestas, revisiones de los procesos y otras documentaciones.

- Establecer una línea base, para medir el impacto que tiene la implementación de la solución.

4.2.6 Estructura de coste

Los costos relacionados al desarrollo e implementación del servicio descrito son los siguientes:

- Marketing: Creación y Mantenimiento de los medios digitales y físicos que sirven para propagar la voz de la empresa y los servicios que ofrece.
- Viáticos: para los movimientos de personal que se dan en los momentos que se realiza encuentros con clientes potenciales.
- Recursos Humanos: SUMADOCASI, está constituido por el momento por 5 integrantes los cuales se benefician en partes iguales de acuerdo a la utilidad generada.

4.2.7 Fuentes de ingresos

Este proyecto en una primera etapa está orientado a ser una implementación dentro de una empresa por lo que no se contempla una fuente de ingreso perse. La empresa estaría financiada el mismo debido a que el proyecto representa los siguientes beneficios que impacta sus términos financieros:

- Mayor ahorro monetario por control de hardware: mejorando las frecuencias de mantenimientos, previniendo fallos que impacte las operaciones y eficientizando la vida útil de los equipos obtendremos un beneficio tanto a corto como a largo plazo monetariamente.
- Crecimiento de clientes de nuestro usuario: el compromiso que genera nuestro usuario a sus clientes con una experiencia de estabilidad en sus redes permite ser recomendado a muchos más clientes de manera orgánica.

4.2.8 Misión, Visión y Objetivos

Misión

Proveer servicios de consultoría y capacitación de servicios de inteligencia de negocios y Big Data en el sector de las telecomunicaciones, proporcionando soluciones que permitan mejorar el sistema de servicios que ofrecen.

Visión

Convertir a SUMADOCASI en la respuesta a las necesidades de mejoras que las empresas de telecomunicaciones demandan, conectando su negocio con los beneficios que ofrecen los servicios de BI y Big Data.

Objetivos

SUMADOCASI tiene el objetivo de convertir el sector de las telecomunicaciones en la República Dominicana en empresas dirigidas por los datos (data driven companies) con el uso de herramientas de Machine Learning y Big Data. En general se pretende proveer soluciones tecnológicas para brindar visibilidad sobre el comportamiento de las infraestructuras utilizadas para proveer los servicios de telecomunicaciones que se ofrecen hoy en día en la República Dominicana, para ello se estarían utilizando los propios recursos con los que cuenten dichas empresas, principalmente los datos arrojados por cada elemento que compongan dicha infraestructura.

Los principales objetivos que agregan valor al negocio de las telcom:

- Incremento en la satisfacción de los clientes con el servicio
- Incremento en los ingresos generales de las empresas
- Incremento en la comprensión del negocio, data driven business
- Reducción de costos operativos en el proceso de solución de averías
- Reducción en la tasa de deserción de clientes
- Reducción en la tasa de mantenimiento reactivo-correctivo
- Reducción en la tasa de reportes de averías de clientes
- Eficientizar en el uso de recursos empleados en el mantenimiento preventivo
- Redefinición de las estrategias empleadas en el mantenimiento preventivo

Capítulo 5 - Plan de acción

5 - Plan de acción

5.1 Alcance del proyecto

Para el desarrollo y correcta implementación de este proyecto de BI/BA y Big Data, el insumo crítico son los datos. La intención es crear una sinergia automatizada entre las herramientas usadas para la minería de datos y los datos disponibles a utilizar que permita a WIND Telecom identificar posibles problemas que afecten la satisfacción de sus clientes en los servicios que ofrecen sobre su red LTE. Con esto se pretende que WIND Telecom mejore su gestión de recursos y planes estratégicos en la ejecución de mantenimientos preventivos, traduciéndose en un aumento de la rentabilidad y calidad del servicio ofertado.

En una primera fase es necesario obtener los datos que se utilizarán como insumo al sistema, estos datos se obtendrán de dos fuentes fundamentales: sistemas automáticos que graban los datos progresivamente y datos creados de forma manual. Los datos históricos permitirán definir los modelos predictivos necesarios para estimar las posibles averías de los elementos que componen la red LTE de WIND Telecom. Todo el procesamiento de los datos será desarrollado in-house, utilizando la infraestructura TI de WIND Telecom.

Para luego ofrecer la facilidad de poder acceder a toda la información resultante en tiempo real desde cualquier dispositivo o lugar, utilizando para este fin Power BI como herramienta de visualización de datos.

5.1.1 Objetivos específicos

- Incremento en la satisfacción de los clientes con el servicio
- Incremento en los ingresos anuales de la empresa
- Incremento en la comprensión del negocio, data driven business
- Reducción de costes operativos
- Reducción en la tasa de deserción de clientes, por voluntad propia
- Reducción en la tasa de mantenimiento reactivo-correctivo
- Reducción en la tasa de reportes de averías de clientes

- Eficientización en el uso de recursos humanos y económicos empleados en el mantenimiento preventivo
- Redefinición de las estrategias empleadas en el mantenimiento preventivo

5.1.2 Métricas

Con la finalidad de cuantificar de forma efectiva el avance hacia los objetivos planteados, gracias a la implementación de la solución planteada en este informe, se estará utilizando un conjunto Indicadores Claves de Gestión (KPIs en sus siglas en ingles). Este conjunto de indicadores es a continuación nombrado en el cuadro mostrado debajo.

TABLA INDICADORES CLAVES WIND TELECOM (KPIs)								
No	**Indicador**	**Desc.**	**¿Quién?**	**¿Con qué?**	**¿Cómo?**	**¿Para qué?**	**Frec.**	**¿Cuánto?**
1	**Ingresos Neto**	Ingreso neto anual	Patricio Ferreiras	DB CRM	Ingreso neto por año	Medir rentabilidad negocio	Anual	Incremento 5% anual
2	**Costes Operativos**	Total costo operativo red LTE	Elvy Duran	Presupuesto operacional	Presupuesto Mantenimiento de la Red	Observar si la solución es efectiva	Anual	Reducción del 30% en primera fase
3	**Churn Rate (%)**	Tasa deserción clientes red LTE	Jose Raúl Pichardo	DB CRM	Total Desconexión / Total de suscripciones	Medir la tasa de deserción de cliente	Mensual	Mantener por debajo del 5%
4	**Mantenimiento Correctivo (%)**	Tasa de acciones Mant. correctivo	Rafael Pichardo	Service Manager	Mant. correctivos averías / Total Averías	Medir la eficacia de modelo.	Mensual	Menos del 60%
5	**Averías clientes LTE (%)**	Tasa de averías clientes red LTE	Elvy Duran	DB CRM	Reportes de averías clientes / Total de suscripciones	Para medir la calidad del servicio	Semanal	Menos del 7%

Tabla 2. *Indicadores claves de WIND Telecom*

5.2 Análisis de actividades y tareas

5.2.1 Modelo lógico

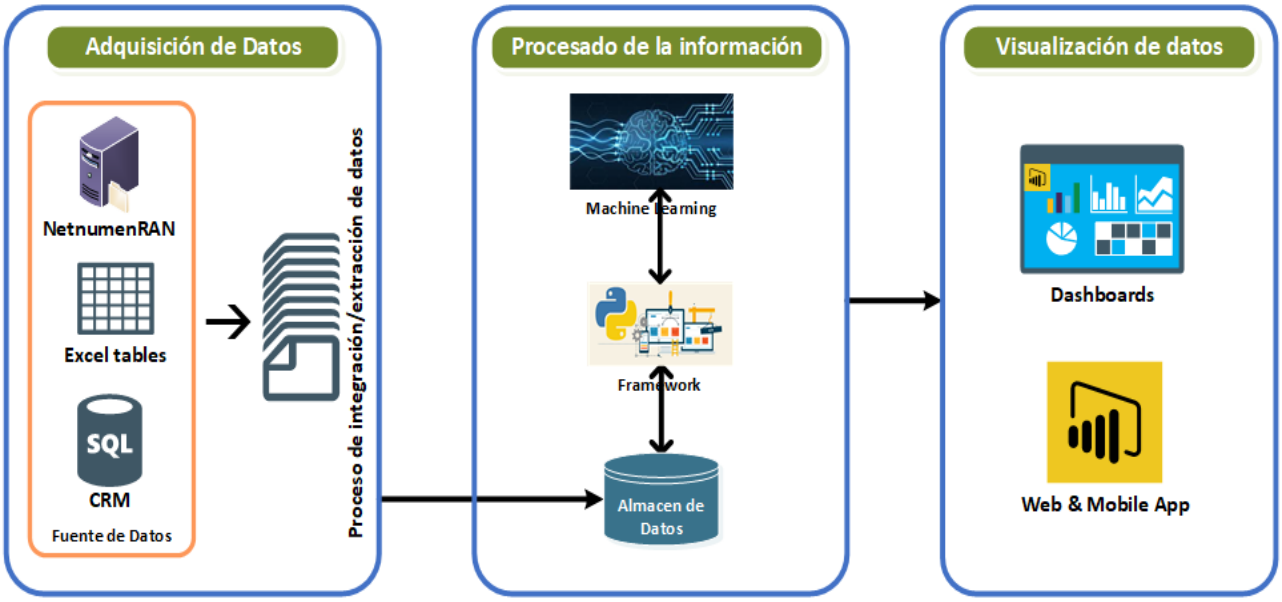


Imagen 8. *Modelo Lógico de la Solución Propuesta.*

5.2.2 Fase desarrollo de la solución propuesta

I. Identificación de los datos

En esta fase se identificaron los datos elegidos como los de interés para la ejecución de la solución propuesta. Los mismos fueron segmentados en dos grandes grupos según la utilidad de cada dato, ya que algunos se estarán utilizando para medir el progreso y otros para realizar el modelado. Debajo se presentan el listado de estas informaciones y una breve descripción de ellas. para luego mostrar un cuadro donde se muestra de forma resumida el detalle de las informaciones.

Datos para medir el progreso:

- **Churn Rate Red LTE.** Este consiste en el resultado de dividir el total de clientes desconectados, por el total de suscripciones de clientes activos. El mismo es generados en el dpto. de Operaciones de Servicios.

- **Encuestas Satisfacción Clientes Red LTE.** son las encuestas realizadas al cliente, se mide mensualmente de forma manual. Estos datos son generados en el dpto. Operaciones de Servicios.
- **Costos Operacionales Red LTE.** Este consiste en la distribución de los costes operativos generales a través de todos los departamentos y subdepartamentos. Estas informaciones actualmente se manejan en una hoja de Excel y se actualizan de forma anual a través de la planificación de presupuesto empresarial. Estos datos son generados en el dpto. Tecnología y Operaciones de la Red.
- **Entrada Averías celdas Red LTE**. Se generan cuando el operador del Network Operation Center visualiza alarmas de averías y las refiere al dpto. de mantenimiento para su solución. Se utiliza el HPE Service Manager como gestor de problemas o incidentes. Estos datos son generados en el Network Operation Center.
- **Disponibilidad celdas Red LTE**. Es el resultado de la división del total del tiempo que la celda se mantuvo activa entre el tiempo total del periodo de medición. Actualmente existen dos formas de generarlos, con los datos de Entrada de Averías Red LTE y con los datos generados por el Netnumen RAN ("Cell Availability"). Estos datos son generados en el Network Operation Center.

Datos para modelado:

- **Alarm History Red LTE.** Este consiste en el histórico de alarmas registradas ante las averías presentadas. Es la fuente de datos que se estará utilizando para la creación de los modelos, y así poder predecir averías o fallas en el hardware de las BBUs. Estos datos son generados en el Netnumen RAN, plataforma que monitorea en tiempo real el estado de los Network Equipment (NE) o BBUs para el caso de interés.

Todas estas informaciones serán debidamente recolectadas, transformadas para la maximización de su utilización y almacenadas en una base de datos para su posterior explotación.

Dato	Tipo	Uso	Origen	Descripción
Churn Rate Red LTE	Estructurado	Medir progreso	CRM DB	Deserción de clientes de la red LTE
Encuestas Satisfacción Clientes Red LTE	Estructurado	Medir progreso	Excel files	Satisfacción de clientes con el servicio brindado sobre red LTE
Costos Operacionales Red LTE	Estructurado	Medir progreso	Excel files	Costos operativos red LTE
Averías Red LTE	Estructurado	Medir progreso y Modelado	Tabla CSV del HPE Service Manage	Datos de incidentes presentados en las celdas de la red LTE
Disponibilidad Red LTE	Estructurado	Medir progreso	NetnumenRAN	Tiempo total celdas en estado activo del periodo medido vs tiempo del periodo medido
Alarm History Red LTE	Estructurado	Modelado	NetnumenRAN	Histórico de alarmas sobre la salud de los elementos del RAN de la red LTE
QoP Red LTE	Estructurado	Medir progreso	NetnumenRAN	Métrica ponderada de los KPIs usados para medir el Performance de la Red LTE

Tabla 3. *Datos identificados como relevantes para el desarrollo de la propuesta.*

II. Captura de datos

En esta fase procederemos a usar las funcionalidades de SQL Server Integration Services (SSIS) para la captura y preprocesamiento de los datos.

1. Proceso ETL para cargar los datos de "*Alarm History Red LTE*", extraído desde el NetnumenRAN, a un repositorio en base de datos.
2. Proceso ETL para extracción de "Desconexiones Total" y "Suscripciones Activas Total" por año-mes para calcular el "*Churn Rate*" y almacenarlos en un repositorio en base de datos.
3. Proceso ETL para carga de los datos de "Encuestas de *Satisfacción de Clientes Red LTE*" desde Excel a un repositorio en base de datos.

4. Proceso ETL para gar de los datos de "Costos Operacionales Red LTE" desde Excel a un repositorio en base de datos.
5. Proceso ETL para consumir los datos de "*Averías de la Red*" desde la tabla "SMG.NetworkIncidents", ya existe un repositorio en la empresa con la información estructurada.
6. Proceso ETL para consumir datos de "*Disponibilidad de la Red*" desde la tabla "SMG.NetworkAvailability", ya existe un repositorio en la empresa con la información estructurada.
7. Proceso ETL para consumir datos de "QoP Red LTE" desde la tabla "*LTE.RAN_PerformanceMonthly_QoP*", ya existe un repositorio en la empresa con la información estructurada

III. Almacenamiento de los datos

Se creará una estructura de datos usando tecnología SQL, específicamente el motor de base de datos SQL Server de Microsoft. El mismo contará con tres tipos de repositorios:

1. Repositorio para datos crudos extraídos de las fuentes de datos
2. Repositorio para datos preprocesados en la fase de "*Transformación y validación de datos*"
3. Repositorio para datos arrojados por la fase "*Descubrimiento y Modelado*"

IV. Transformación y validación de los datos

- Perfilado: para el modelo predictivo se utilizaron los datos de Alarmas y Averías con un periodo de información desde noviembre 2018 a septiembre 2019. Las informaciones de las Alarmas se agrupan en 5 tipos, la cuales son: Communications Alarm, Environmental Alarm, Equipment Alarm, Processing Error Alarm y Quality of Service Alarm. Dentro de estas 5 mencionadas categorías se encuentra distribuidas un total de 88 alarmas, siendo cada uno parte de un elemento que conforma la Red LTE RAN.
- Limpieza: las informaciones de alarmas y averías que se obtuvieron fueron tratadas través de un proceso de análisis exploratorio de datos (EDA). En este proceso se utilizaron las alarmas como variables principales, y como resultado de observar las 88

variables, cada una de acuerdo con su categoría, obtuvimos cuales son las 37 variables que causarían el mayor impacto en el modelo predictivo.

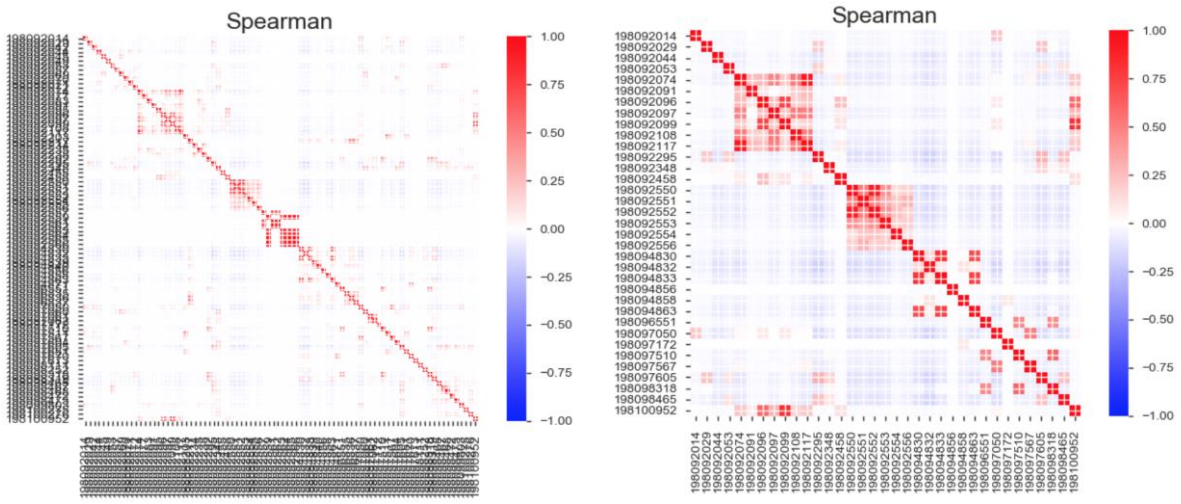


Imagen 9. *Matriz de correlación de las 88 (Izquierda) y 37 (Derecha) alarmas*

- Estructuración: Se estructuraron las informaciones resultantes del proceso de mapeo y el EDA, con la finalidad de que este sea consumido como dataset para el modelo predictivo.

V. Descubrimiento y modelado

El modelo predictivo plantea obtener una lista de los elementos de la LTE RAN, que posean las mayores probabilidades de tener equipos con posibles futuras averías, obteniendo la misma por medios de patrones de alarmas que categorizaran el elemento con un nivel o probabilidad de riesgo de avería alta, al mismo tiempo con un % de supervivencia o funcionamiento de vida de esta.

Para lograr los objetivos anteriores se utilizaron dos tipos de modelo supervisado LogisticRegression (Clasificación) para clasificar si está en avería y Lifelines de KaplanMeierFitter (Survival Analysis) para estudiar mediante el tiempo de vida del elemento la probabilidad de supervivencia del mismo.

Para los procesos del modelado de clasificación podemos destacar los siguientes pasos:

- Entrenamiento: para la selección de los diferentes modelos a trabajar se dividió el dataset en dos subconjuntos de datos, uno para entrenamiento (80 %) y otro para realizar test (20 %). En el entrenamiento se utilizaron varios modelos para detectar cual podría tener mayor precisión en la identificación de averías;

```
# Librerias de manipulacion de datos
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Librerias de modelados
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.neighbors import KNeighborsClassifier

# Librerias de analisis
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
```

Imagen 10. *Librerías utilizadas en el modelo predictivo*

- Predicción: se revisaron las precisiones de los modelos con los resultados obtenidos después de la predicción y los resultados reales de los datos de test, dando la misma para los modelos: LogisticRegression con un 99.62 %, SVC con un 99.75 % y KNeighborsClassifier con un 90.23 %;

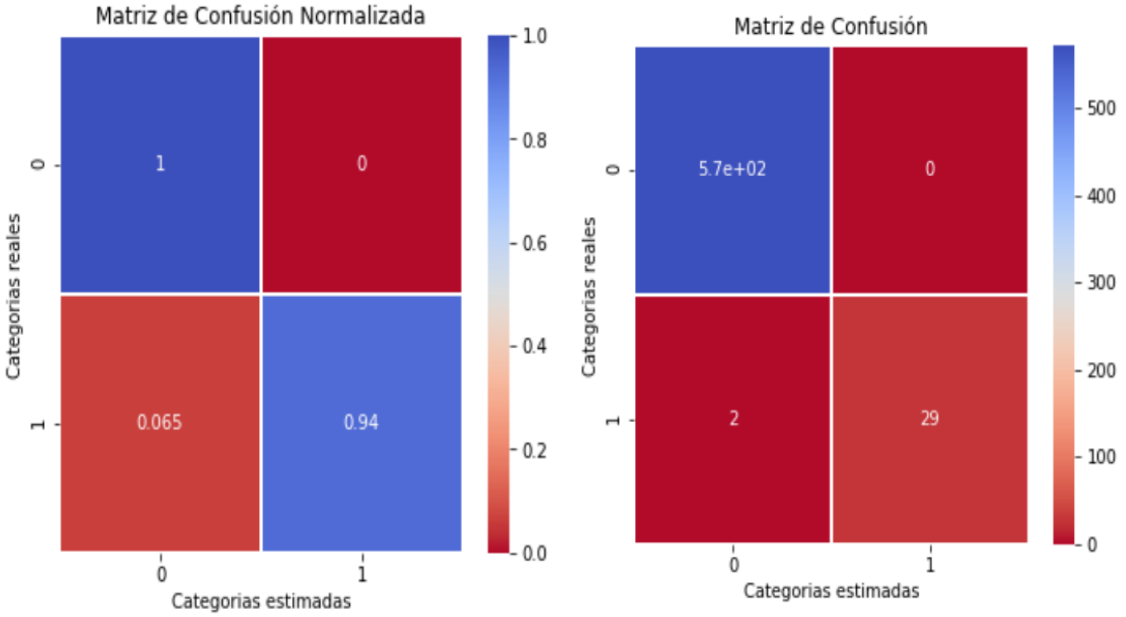
```
# Implementando modelo de de clasificacion LogisticRegression
logreg = LogisticRegression()
logreg.fit(X_train, y_train)
Y_pred = logreg.predict(X_test)
print('Precisión Regresión Logística:')
print(logreg.score(X_train, y_train))

mostrar_matriz_confusion(y_test, Y_pred)
```

```
//anaconda3/lib/python3.7/site-packages/sklearn/linear_model/logis
hanged to 'lbfgs' in 0.22. Specify a solver to silence this warnin
  FutureWarning)

Precisión Regresión Logística:
0.9962655601659751
```

Imagen 11. *Código para el entrenamiento del modelo de Logistic Regression*

- Validación: se utilizó la matriz de confusión, obteniendo las métricas de precisión, recall y f1-score la cual nos permite estudiar la cantidad de acierto que nuestro modelo realiza correctamente. El modelo seleccionado es LogisticRegression con una precisión

de 100 % indicando qué porcentaje de lo que se predijo fue correcto, recall de 94 % indicando que todos los equipos averiados fueron identificados correctamente y f1-score de 97 % siendo esta un promedio de la precision y recall.



	precision	recall	f1-score	support
0	1.00	1.00	1.00	572
1	1.00	0.94	0.97	31
accuracy			1.00	603
macro avg	1.00	0.97	0.98	603
weighted avg	1.00	1.00	1.00	603

Imagen 12. *Matriz de confusión del modelo utilizando LogisticRegression*

Para los procesos del modelado de supervivencia podemos destacar los siguientes pasos:

- Entrenamiento: para el entrenamiento de este modelo se construyó un dataset con las duraciones de funcionamiento activo de los elementos de cada BBU con un periodo de información del 2014 al 2019;

```
BBUTypes = ["SA", "BBU", "BPL1",
            "CCC", "FCE", "R8978 S2600L",
            "BPN2", "CCE1"]

KmfBT  = {}

# Implementando modelo de supervivencia
for bt in BBUTypes:

    print(bt)
    df_selected = df_train[df_train.BoardType == bt]

    kmf = KaplanMeierFitter()
    kmf.fit(durations = df_selected.Duracion, event_observed = df_selected.Averias)

    predict = kmf.predict([2,3,5,6,10])

    KmfBT[bt] = kmf

    print(predict)
    print('')
```

```
SA
2     0.956522
3     0.751553
5     0.318841
6     0.000000
10    0.000000
```

***Imagen 13**. Código para el entrenamiento del modelo lifelines de KaplanMeierFitter*

- Salida de información: Como resultados se obtuvo que las mayorías de los elementos tienen una probabilidad muy baja de funcionamiento en un promedio de vida de 4 a 6 años;

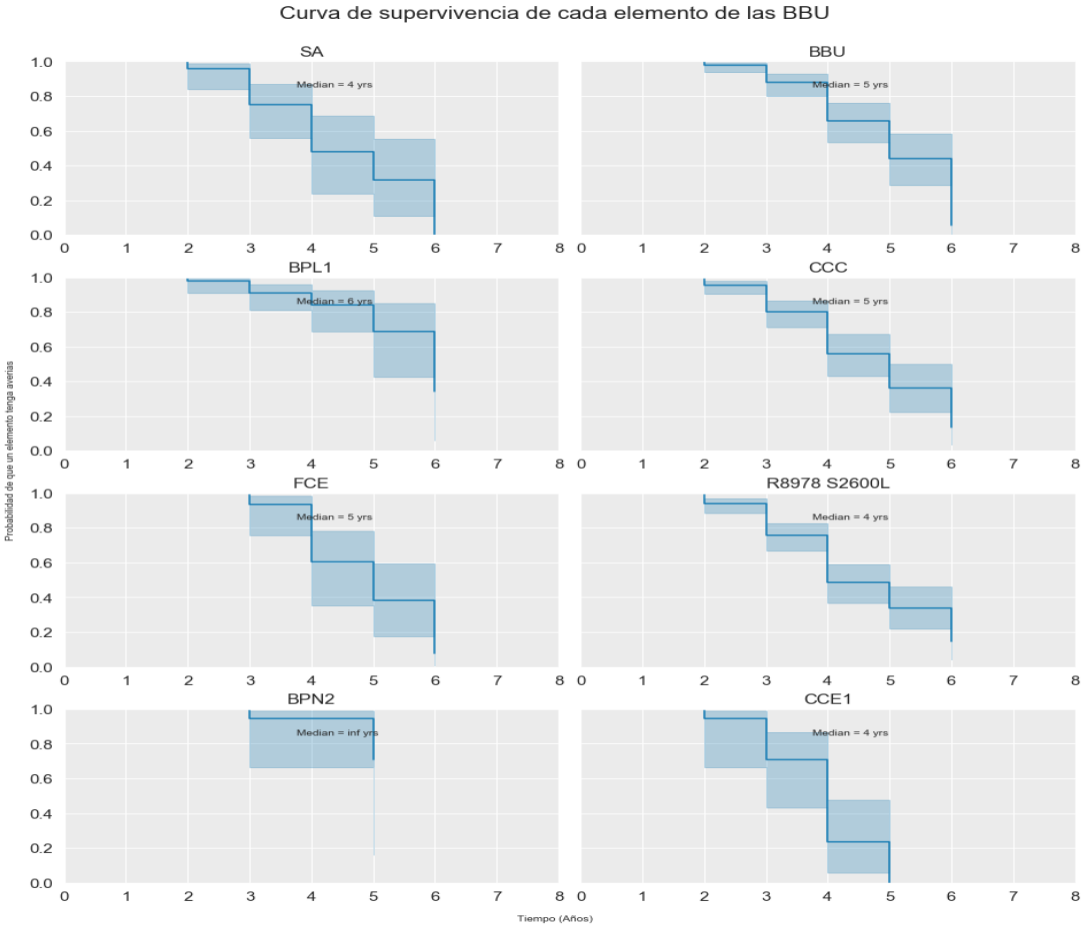


***Imagen 14**. Curva de supervivencia de elementos de una BBU con el modelo lifelines de KaplanMeierFitter*

VI. Visualización

Los resultados finales se presentarán en cuadros de mando diseñados sobre la plataforma de MS Power BI ya que WIND Telecom consta con licenciamiento para la misma. Esta herramienta permite que los consumidores de la información puedan acceder y visualizar los informes a través de los servicios web o móvil de MS Power BI desde cualquier parte del mundo y a cualquier hora deseada. La gran ventaja que esta herramienta ofrece es que tanto el equipo ejecutivo como operativo tendrán acceso rápido y remoto.

El objetivo principal es poder mostrar la evolución de los datos con la granularidad necesaria para WIND Telecom. Para ello procederemos inicialmente a dividir las informaciones a visualizar en dos grandes bloques:

Resumen Ejecutivo:

Este cuadro de mando será utilizado para mostrar los resultados de las métricas objetivos de interés al nivel ejecutivo o estratégico de WIND Telecom, sirviendo como herramienta de observación continua para 3 de los KPIs que actualmente se les da un seguimiento continuo en la empresa. Adicionalmente les ayudará a medir la eficacia de la solución implementada.

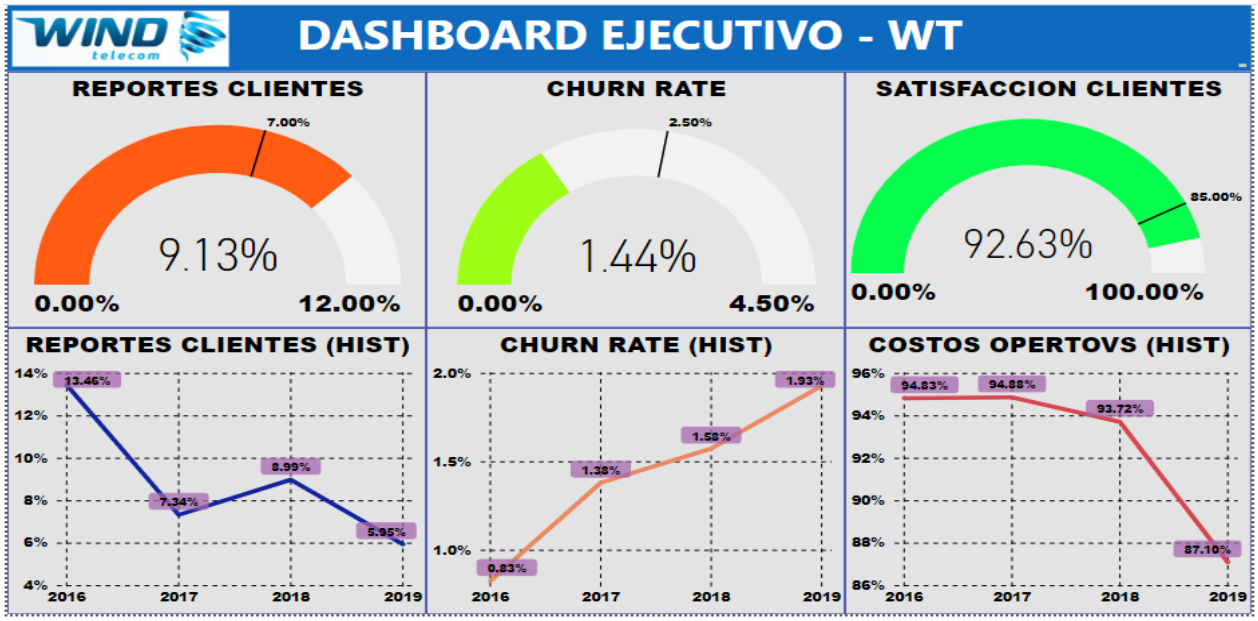


Imagen 15. *Dashboard Ejecutivo. Muestra tres de los indicadores estratégicos.*

Resumen Operativo:

Este cuadro de mandos será utilizado para mostrar los resultados de las métricas objetivos de interés al nivel operativo en WIND Telecom. También se creará un cuadro de mandos para ver los resultados arrojados por el modelo.

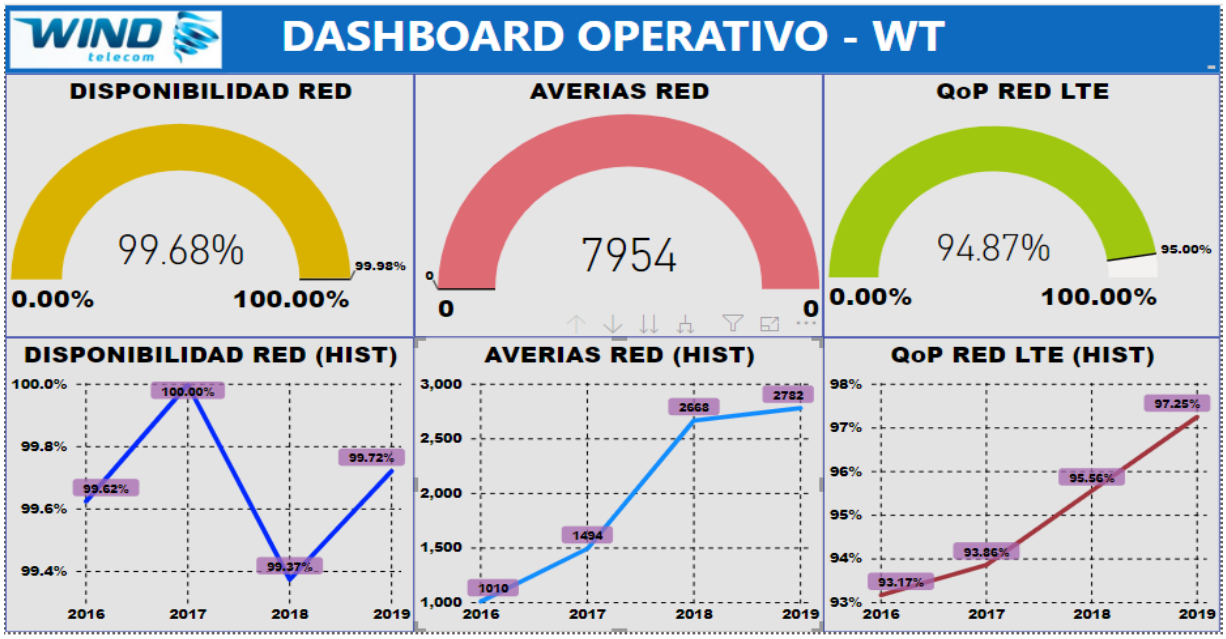


Imagen 16. *Dashboard Operativo. Muestra tres de los principales indicadores operativos*

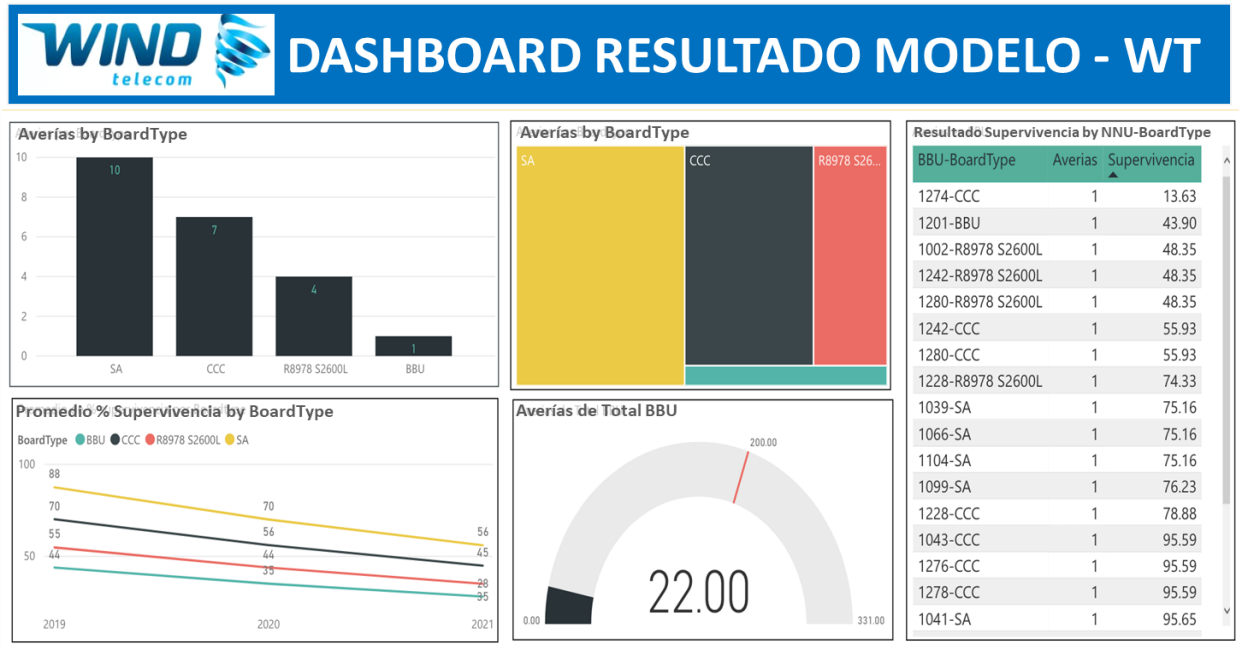


Imagen 17. *Dashboard Operativo. Muestra alarmas por board type, un acumulado y la tabla de supervivencia por BBU-BoardType*

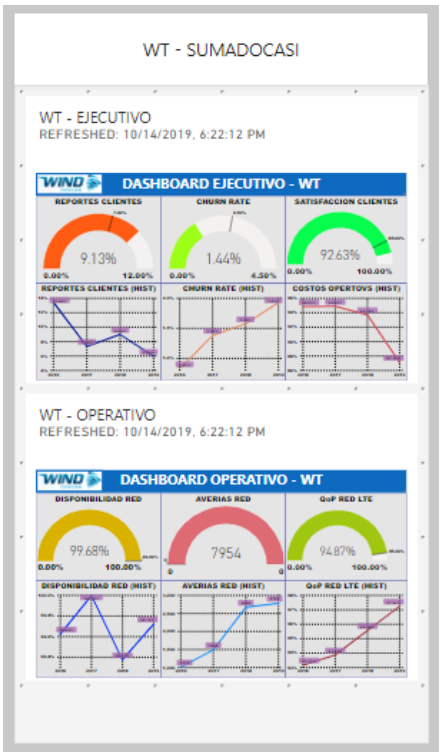


Imagen 18. *Dashboard Solución Visual Propuesta (Visión Móvil).*

5.2.3 Arquitectura técnica de la solución propuesta

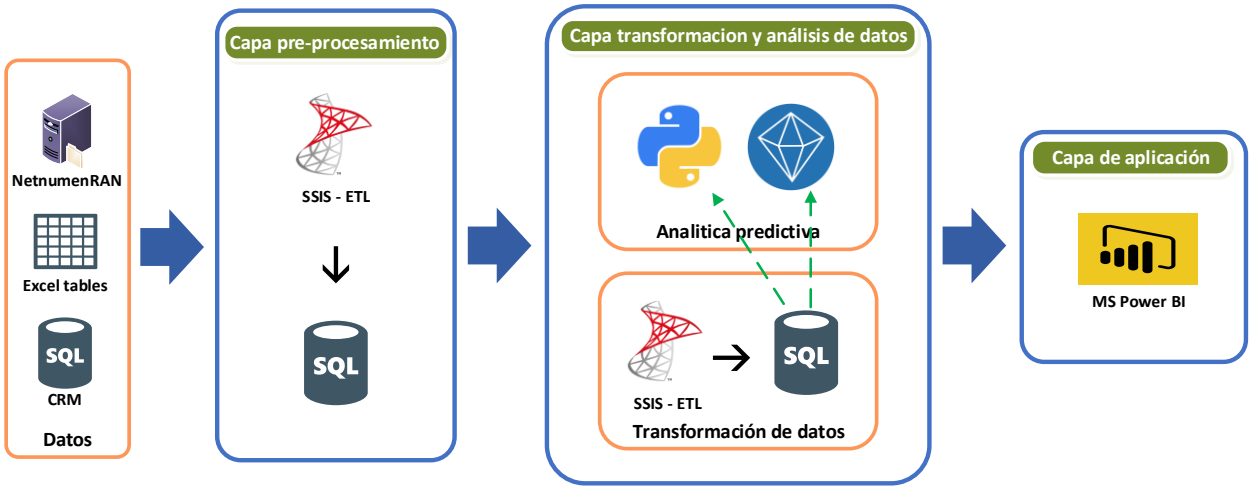


Imagen 19. *Arquitectura técnica de la solución Propuesta.*

5.3 Análisis de recursos

5.3.1 Talento humano

El talento humano, para el éxito de este proyecto juega un papel fundamental, pues es un factor crítico de alto impacto sobre los resultados a obtener. Por tal razón se presta especial y detallada atención a la estructura organizacional y las funciones desempeñadas por cada uno de los involucrados.

Se plantea agregar dos puestos nuevos con la labor de monitorear y analizar la herramienta de predicción. Las otras posiciones ya están creadas, pero a continuación se dará una breve descripción de su participación en el proyecto.

Estructura organizativa

Imagen 20. *Estructura organizacional de WIND Telecom (Solo considerando las áreas impactadas por el problema y la solución planteadas).*

Denominación del recurso humano y tareas asignadas.

Responsabilidades actuales del personal mencionado en el diagrama con sus nuevas funciones anexas en la descripción.

- **CEO**: Es quien dirige a la empresa y la guía entorno al éxito y mayor productividad en el campo de las telecomunicaciones. Este es quien aprobará las sugerencias dadas por el director de datos.
- **VP Tecnología y Operaciones de la Red**: Responsable de los lineamientos tecnológicos, estrategias, proyectos de expansión y operaciones y servicio de la red de telecomunicaciones.
- **Dirección de Operaciones de la Red**: Responsable del esquema de red, mantenimiento general a la red. Su función principal es dar mantenimiento correctivo y preventivo a la red.
- **Gerencia Mantenimiento de la Red:** Responsable de resolver las averías en la empresa, mantener los servicios disponibles y la operabilidad de los equipos de la red.
- **Gerencia Ingeniería y Optimización de Red:** Responsable de la instalación y puesta en marcha de las nuevas localidades, supervisar los equipos LTE y móviles y brindar soporte a los departamentos de mantenimiento y operaciones en la red LTE.
- **Dirección de Datos y Analítica:** Responsable de gestionar los activos informacionales de la empresa, definir estrategias para beneficiar a la empresa con los datos, hacer propuestas para el crecimiento y éxito de la empresa tomando en cuenta los datos, manejar personal analítico de la empresa y llevar a la empresa a tomar las decisiones en base a los datos. El director de Datos es quien canaliza la información sacada del software a la junta directiva con las mejores conclusiones obtenidas del software.
- **Analistas de Datos:** Responsable de analizar, limpiar, inspeccionar y transformar los datos para resaltar informaciones y llevar a la empresa a tomar decisiones en base a los datos. Este analista es quien operará la herramienta de predicción implementada.

Cargas de trabajo

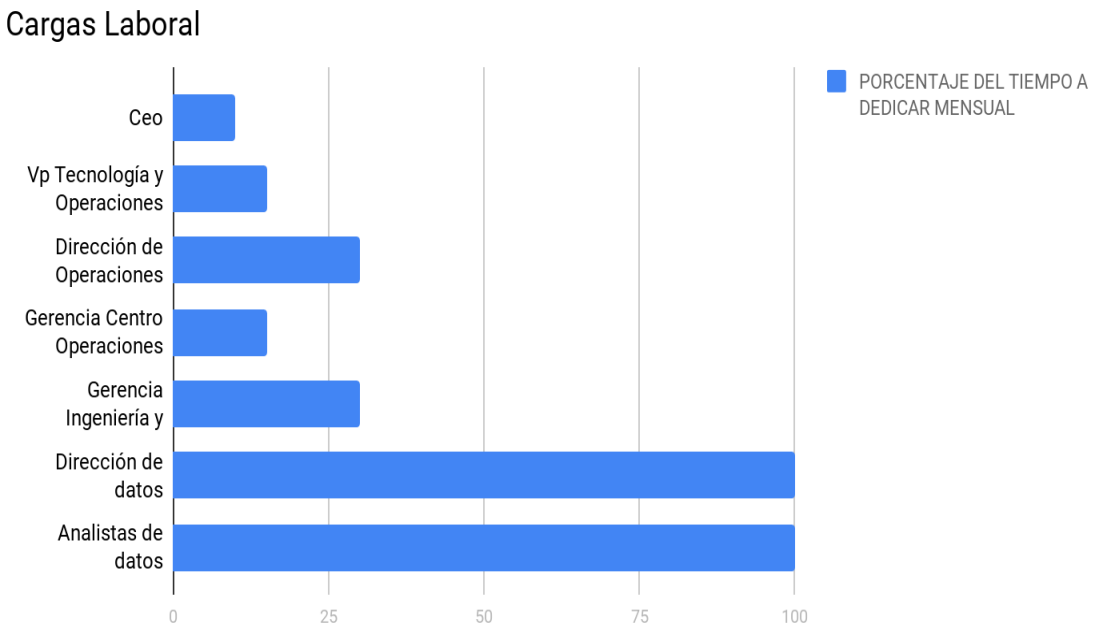


Imagen 21. *Carga Laboral distribuida entre las distintas posiciones de la estructura organizacional considerada.*

El costo de participación

A continuación, se describe el costo de participación de las posiciones laborales requeridas para implementar el proyecto, cabe destacar que solo se tomará en cuenta las posiciones de Dirección de datos y Analistas de datos porque son las nuevas posiciones.

Posición	Director de datos	Analistas de datos	Inversión total
Inversión Mensual	RD$150,000	RD$70,000	RD$220,000
Inversión Anual	RD$1,950,000	RD$910,000	RD$2,860,000

Tabla 4. *Inversión adicional realizada en capital humano.*

5.3.2 Recursos físicos

Infraestructura física (recursos físicos)

Para la puesta en marcha del proyecto WT - SUMADOCASI, la empresa WIND Telecom pondrá a disposición todos los recursos tecnológicos necesarios, ya que la empresa cuenta con una infraestructura tecnológica de punta.

Cabe destacar que en esta primera etapa del proyecto no es necesario utilizar servicios en la nube externos, ya que la empresa cuenta con su propia nube, la misma que pondrá a disposición para la puesta en marcha del proyecto.

Suministros y servicios externos

Los suministros necesarios para la puesta en marcha del proyecto son:

- Espacio en la nube para almacenar, analizar y procesar los datos a tratar.
- Conectividad a la red interna de WIND Telecom y conectividad ilimitada fuera de la localidad.
- Computadoras estacionarias y portátiles eficientes para procesar datos a mediana y gran escala.
- Transporte, alimentos y pago de hora extras a la hora de implementar el proyecto.

Los servicios externos que utilizaremos son:

- Comprar dominio sumadocasi.com.do para visualizar los datos.
- Desarrollar el sitio web sumadocasi.com.do.

5.4 Gestión del tiempo (cronograma)

A continuación, definimos el cronograma del proyecto en función de las actividades y fases identificadas. Para el desarrollo establecemos una duración media de 76 días. Esta duración podrá variar en función de las necesidades específicas del cliente.

5.4.1 Listado de tareas

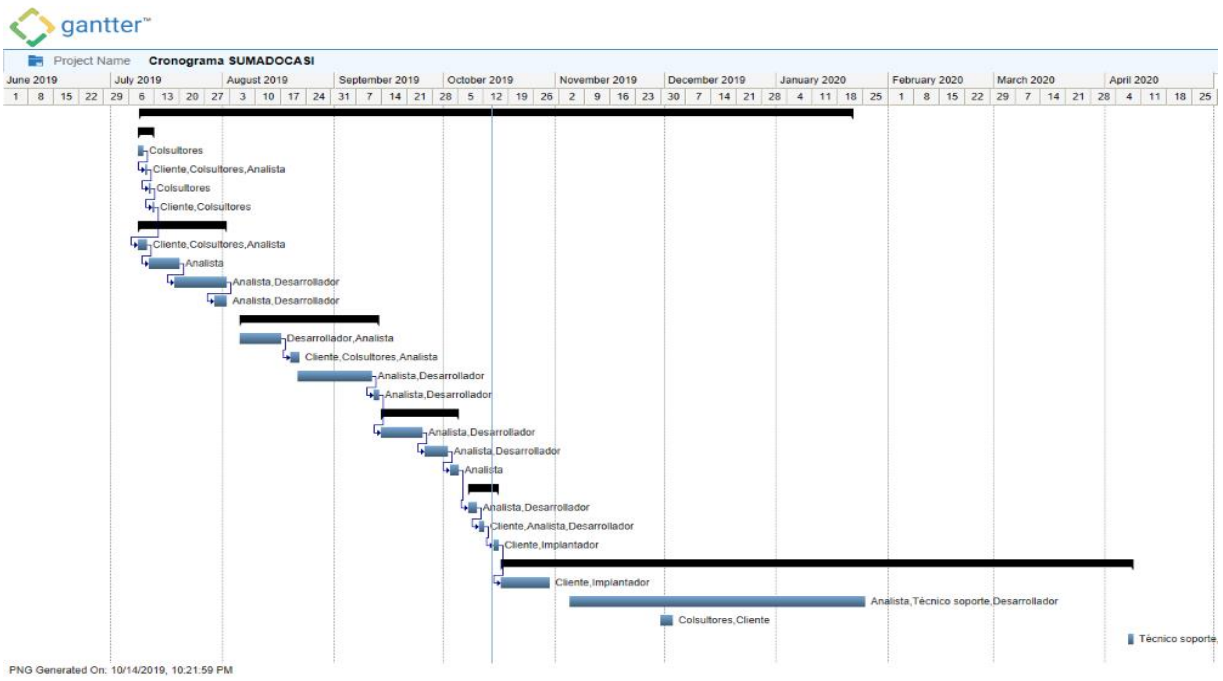
gantter™

Project Name **Cronograma SUMADOCASI**

	Name	Duration	Start	Finish	Predecessors	Resources
1	PROYECTO SUMADOCASI	197days	07/08/2019	01/21/2020		
2	1. DEFINICIÓN ANTEPROYECTO	5days	07/08/2019	07/12/2019		
3	1.1 Definir alcance del proyecto	2days	07/08/2019	07/09/2019		Colsultores
4	1.2 Levantamiento de requerimientos	1day	07/10/2019	07/10/2019	3	Cliente,Colsultores,Analista
5	1.3. Definición de gestión del proyecto	1day	07/11/2019	07/11/2019	4	Colsultores
6	1.4 Presentacion y aprobacion del anteproyecto	1day	07/12/2019	07/12/2019	5	Cliente,Colsultores
7	2. ANÀLISIS	19days	07/08/2019	08/01/2019		
8	2.1 Estudio de requerimientos	3days	07/08/2019	07/10/2019	6	Cliente,Colsultores,Analista
9	2.2 Análisis fuentes de datos	7days	07/11/2019	07/19/2019	8	Analista
10	2.3 Análisis modelos predictivos	11days	07/18/2019	08/01/2019	9	Analista,Desarrollador
11	2.4 Estudio alternativas tecnológicas	4days	07/29/2019	08/01/2019	10	Analista,Desarrollador
12	3. DISEÑO	29days	08/05/2019	09/12/2019		
13	3.1 Selección modelo análisis	10days	08/05/2019	08/16/2019		Desarrollador,Analista
14	3.2 Definición indicadores de resultado	3days	08/19/2019	08/21/2019	13	Cliente,Colsultores,Analista
15	3.3 Definición de arquitectura técnica	15days	08/21/2019	09/10/2019		Analista,Desarrollador
16	3.4 Solicitud de recursos tecnológicos	2days	09/11/2019	09/12/2019	15	Analista,Desarrollador
17	4. IMPLEMENTACIÓN	16days	09/13/2019	10/04/2019		
18	4.1 Desarrollo de modelos predictivos	8days	09/13/2019	09/24/2019	16	Analista,Desarrollador
19	4.2 Integración solución técnica	5days	09/25/2019	10/01/2019	18	Analista,Desarrollador
20	4.3 Desarrollo de cuadro de mandos	3days	10/02/2019	10/04/2019	19	Analista
21	5. VALIDACIÓN Y DESPLIEGUE	7days	10/07/2019	10/15/2019		
22	5.1 Prueba del modelo predictivo	3days	10/07/2019	10/09/2019	20	Analista,Desarrollador
23	5.2 Pruebas del sistema	2days	10/10/2019	10/11/2019	22	Cliente,Analista,Desarrollador
24	5.3 Puesta en marcha	2days	10/14/2019	10/15/2019	23	Cliente,Implantador
25	6. SOPORTE Y MANTENIMIENTO	125days	10/16/2019	04/07/2020		
26	6.1 Formar a los usuarios	10days	10/16/2019	10/29/2019	24	Cliente,Implantador
27	6.2 Soporte a los usuarios	60days	11/04/2019	01/24/2020		Analista,Técnico soporte,Desarrollador
28	6.3 Seguimiento	2days	11/29/2019	12/02/2019		Colsultores,Cliente
29	6.4 Mantenimiento	2days	04/06/2020	04/07/2020		Técnico soporte,Analista,Desarrollador

PNG Generated On: 10/14/2019, 10:16:27 PM

5.4.2 Diagrama de Gantt



Capítulo 6 - Optimización de los resultados

6 - Optimización de los resultados

6.1 Proyecto de optimización

Hoy en día las empresas en su necesidad de poder coexistir en un mercado altamente competitivo necesitan mantenerse en todo momento actualizados en los aspectos claves que le requiere el mercado en el cual participa en particular. En el caso de las empresas de Telecomunicaciones sus necesidades principalmente se basan en adquisición de tecnología de punta, disminución de costos operativos y aumento en la percepción positiva de sus clientes.

En vista de las necesidades planteadas SUMADOCASI le presenta a WIND Telecom, empresa de telecomunicaciones, un gran proyecto de optimización de recursos internos, los cuales se verán reflejados directamente en la satisfacción de los clientes de la empresa de telecomunicaciones, hacia una experiencia más positiva. Dicho proyecto se basa en la implementación en un sistema visual predictivo que ayude a la empresa a maximizar los recursos (personas, tiempo, materiales) invertidos en las actividades realizadas en el área de mantenimiento, a través de la gestión inteligente de las averías que actualmente aparecen de forma inesperada.

6.1.1 Beneficios tangibles.

Generación de ingresos

La implementación del proyecto de predicción de averías en el sistema de Red LTE de WIND Telecom, impactará de manera positiva tanto directa como indirectamente a toda la empresa, en vista de que la solución planteada ayudará en la creación de mayores ingresos a los percibidos actualmente. Logrando:

- *Incrementar la captación clientes nuevos y recuperación aquellos que anteriormente abandonaron.* Debido al aumento de la *satisfacción* de los clientes que se generará al gestionar las averías inteligentemente, minimizando el perjuicio del servicio al cliente, lo cual creará de forma indirecta una experiencia global (voz pública), aumentando así la *credibilidad.*

- *Incrementar los niveles de éxito en la gestión Ventas realizadas para todos los servicios ofrecidos por WIND Telecom,* al contar con retroalimentación positiva disponible como respaldo.
- *Ampliar las zonas de cobertura, incrementando de forma exponencial la captación de clientes.* Ya que al contar con mayores niveles de credibilidad el público interesado en contratar los servicios de WIND Telecom aumentará, fomentando así la expansión de la red y las capacidades.

Algunos indicadores necesarios para la medición de dichos beneficios son:

- Satisfacción de los Clientes, y su variación
- Ingresos generados por áreas de cobertura.
- Clientes Captados
- Clientes re-afiliados

Reducción de costes

Debido a la naturaleza del servicio ofrecido, los costos de las empresas de telecomunicaciones se centran en su inversión tecnológica, el mantenimiento de su infraestructura tecnológica y en el capital humano que este posee el cual se encuentra altamente capacitado para poder responder ante las eventualidades y necesidades de las empresas de telecomunicaciones.

Como resultado de la implementación de la mejora propuesta, WIND Telecom además de generar mayores ingresos debido a lo ya expuesto en el punto anterior, obtendrá también:

- *Reducción de Costes operativos de mantenimiento los cuales involucran tiempo de personal capacitado,* Debido a que las averías son previstas los equipos de labor que estarán trabajando los mantenimientos preventivos serán planificados y organizados según la criticidad del mantenimiento *(Incremento de la eficiencia Operativa*).
- *Reducción de costos de compras de materiales/partes para reparación los cuales actualmente son adquiridas en urgencia no teniendo así la capacidad de negociar precio de adquisición.*

El seguimiento a la reducción de costes se puede realizar a través del estudio de los siguientes indicadores:

- Tiempo Invertido en el mantenimiento de las averías.
- Cantidad de mantenimientos realizados por mes
- Costo total invertido en recursos para mantenimiento de las piezas.

6.1.2 Beneficios intangibles:

Otros beneficios obtenibles de los que WIND Telecom gozará son:

- *Contar con información utilizable para la toma de decisiones*, ante las operaciones y vicisitudes que atraviesa el área de mantenimiento.
- *Utilizar el proyecto de predicción de averías como piloto*, con la finalidad de implementar proyectos similares en donde los datos se conviertan en información útil, comprensible y accesible para la toma de decisiones de las áreas correspondientes.
- *Fomentar una cultura de una Empresa Orientada en Datos*, en donde el capital humano de la empresa se preocupa por verificar la información existente, hacer registros de calidad y tomar decisiones basada en información más que solo de la experiencia del personal.
- *Mayor control de las operaciones.*
- *Facilidad para la expansión en su cobertura*, al sentirse con la capacidad para responder hábilmente ante las posibles averías.

6.1.3 Beneficios estratégicos:

La mesa directiva de la empresa WIND Telecom también experimentará beneficios adicionales los cuales podrían ser tangibles tanto de inmediato como a mediano plazo. A continuación, se listan algunos:

- *Incremento en el conocimiento de la situación actual de la empresa*, debido a la visibilidad que se estará generando en el área operativo de mantenimiento.
- *Mejora en la apreciación del público cliente ante WIND Telecom*, lo cual le facilita la adecuación oportuna de los precios al que venden los servicios.
- *Atracción de otros inversionistas*, puesto a que el nivel de riesgo de pérdida de inversión decaerá debido a la inversión positiva en tecnología y la experiencia positiva de clientes.

6.2 Análisis de rentabilidad del proyecto

Como en todo proyecto de optimización, el fin que busca cualquier cliente es el de poder mejorar sus ingresos o reducir su costo, y para ello es necesario medir la rentabilidad y el retorno de la inversión de este. Para ello nos valemos de tres métricas fundamentales: **VAN, TIR y ROI.**

En el siguiente cuadro podemos ver los flujos de caja esperado para WIND Telecom una vez implementado el proyecto ofrecido.

ANÁLISIS RENTABILIDAD DEL PROYECTO				
	AÑO 0	**AÑO 1**	**AÑO 2**	**AÑO 3**
INVERSIÓN				
Hardware	$ 80.000,00	$ 30.000,00	$ 25.000,00	$ 20.000,00
Software	$ 150.000,00	$ 50.000,00	$ 20.000,00	$ 15.000,00
TOTAL INVERSIONES	**$ 230.000,00**	**$ 80.000,00**	**$ 45.000,00**	**$ 35.000,00**

	AÑO 0	**AÑO 1**	**AÑO 2**	**AÑO 3**
INGRESOS/BENEFICIOS				
Costo operativo: solución de averías	$ 75.000,00	$ 90.000,00	$ 100.000,00	$ 125.000,00
Mejora de reputación	$ 25.000,00	$ 27.500,00	$ 33.000,00	$ 42.900,00
Mejora de retención de clientes	$ 45.000,00	$ 49.500,00	$ 59.400,00	$ 77.220,00
Reducción de averías	$ 60.000,00	$ 66.000,00	$ 79.200,00	$ 95.000,00
Mejora de la competitividad	$ 40.000,00	$ 44.000,00	$ 52.800,00	$ 68.640,00
TOTAL INGRESOS	**$ 245.000,00**	**$ 277.000,00**	**$ 324.400,00**	**$ 408.760,00**

	AÑO 0	AÑO 1	AÑO 2	AÑO 3
GASTOS				
Consultoría	$ 11.000,00	$ 9.000,00	$ 6.000,00	$ 4.000,00
Viajes y desplazamientos	$ 18.000,00	$ 16.000,00	$ 14.500,00	$ 12.000,00
Impuestos	$ 24.500,00	$ 26.950,00	$ 32.340,00	$ 42.042,00
Formación	$ 7.500,00	$ 7.500,00	$ 5.400,00	$ 3.900,00
Mantenimientos correctivos: solución averías	$ 115.000,00	$ 95.000,00	$ 85.000,00	$ 75.000,00
TOTAL GASTOS	**$ 176.000,00**	**$ 154.450,00**	**$ 143.240,00**	**$ 136.942,00**

FLUJO DE CAJA OPERATIVO	**$ (161.000,00)**	**$ 42.550,00**	**$ 136.160,00**	**$ 236.818,00**
VALOR ACTUAL	**$ (161.000,00)**	**$ 38.681,82**	**$ 112.528,93**	**$ 177.924,87**
ACUMULADO	**$ (161.000,00)**	**$(122.318,18)**	**$ (9.789,26)**	**$ 168.135,61**

VAN:	**$ 490.135,61**
TIR:	**49,20%**
ROI:	**213,10%**
PAY-BACK	**2,04**

Tabla 5. *Análisis rentabilidad proyecto SUMADOCASI – WIND Telecom.*

Como podemos apreciar en la tabla rentabilidad del proyecto es muy buena, casi un 50% y con un pay-back de 2 años aproximadamente, lo que representa un buen caso de negocio para el cliente.

VAN	TIR	PAY-BACK	ROI
$ 490.135,61	**49,20%**	**2.04 años**	**213,10%**

Capítulo 7 - Conclusiones

7. Conclusiones

Actualmente las empresas de telecomunicaciones de República Dominicana están dando un giro hacia la transformación digital, buscan optimizar sus recursos disminuyendo los gastos y elevando las experiencias de sus clientes.

El reto para las TELCO es adaptar las soluciones del mercado a su infraestructura actual, encontrar un sistema que optimice sus recursos prediciendo las fallas que pueden suceder durante su operación era imposible sin la ayuda de la Inteligencia de Negocios y el Big Data.

Dichas fallas técnicas producen pérdidas económicas por medio de la insatisfacción del cliente, multas por el INDOTEL que es el órgano regulador de las empresas de telecomunicaciones en República Dominicana y compras de equipos no programadas que llevan a detener la productividad y expansión de la empresa.

SUMADOCASI es la solución a todos los retos planteados por las empresas de telecomunicaciones, mejorar los sistemas de servicios que ofrecen nuestros clientes es nuestro objetivo principal.

Esta herramienta surgió para mejorar el sistema de servicios ofrecido por la empresa de telecomunicaciones WIND TELECOM a través de técnicas de minería de datos: usando machine learning, tecnologías de base de datos y análisis estadísticos.

Por medio de nuestra solución la empresa WIND TELECOM se posiciona entre una de las mejores empresas de telecomunicación de República Dominicana, esta será una solución productiva y rentable para la empresa ya que su inversión será recuperada en un periodo de 3 años.

Anexos

Anexo A - Resultado de Entrevistas

1. Resultados entrevista a Patricio Ferreiras

¿Nombre? Patricio Ferreiras

¿Área de labor? Área de tecnología y operaciones en la red

¿Tiempo laborando en la empresa? 12 años

¿Cargo desempeña? Vicepresidente de tecnología y operaciones de la red

¿Funciones que desempeña?

A. Responsable de los lineamientos tecnología, la estrategia que siga la empresa en función de su evolución tecnológica
B. Responsable de los proyectos de construcción de expansión de G y también proyecto que está dirigido a cliente y sobre todo de negocio
C. Responsable de la operación y del servicio que se le brinda sobre la red de telecomunicaciones

¿Qué tipo de informaciones manejan en su área?

Indicadores de desempeño del área, información de uso de capacidades, tipo de servicio, niveles de servicios que brindamos y proponemos a los clientes, informaciones que brindamos a otras áreas para el desempeño de sus funciones como por ejemplo proyección de mercado, implementación de producto, planes y estrategia de servicio.

¿Estas informaciones la produce el área o la consumen de otras áreas o sistemas?

Mayormente producimos la información porque son información que se produce en la red o en la interacción con los clientes y la recogemos en plataforma.

Como responsable del área de sistemas de soporte (SVSS) producimos información que no son de la red sino por interacciones con el cliente que son más bien utilizado por otras áreas como Mercadeo, Inteligencia de negocio y Operaciones de servicios.

Consumismo tanto las informaciones que producimos, como las recolectada por otras áreas.

¿Cómo se almacenan las informaciones producida por el área? (físico o digital) (Base de datos o archivos)?

Base de datos y archivos planos

¿Las informaciones de otras áreas o sistemas provienen de Base de datos o archivos?

Base de datos y Excel.

¿Utilizan alguna herramienta para la realización de sus funciones diaria, cual/es?

Ofimática, Herramienta de manejo de BI y Big Data (Pininos)

¿Esas herramientas producen las informaciones consumidas por el área o usan como fuentes esas informaciones producidas por el área?

Si, Tantos las informaciones que nosotros construimos, como las informaciones de otros departamentos y de pequeñas herramientas externas.

¿Sabe usted cómo su área impacta en la calidad del servicio ofertado a los clientes?

Claro, el área de tecnología es un área fundamental en el servicio que se le da a los clientes, este no es un área de tecnología que da soporte al negocio, es la fábrica del negocio y todo el servicio se monta en la misma. La estabilidad de esta fábrica en las redes y del desempeño depende directamente del nivel de servicio que reciben los clientes, cada evento, afección, capacidad que está saturada o desempeño en término de eficiente en ese segmento de la red, se siente en muchos servicios del cliente y todo esto impacta a los indicadores fundamentales de servicios.

¿Sabe usted cómo su área impacta en el proceso de solución de las averías en la red móvil?

En alto porcentaje es responsabilidad de esta área de manera indirecta o directa, cada elemento que falla está bajo responsabilidad total de esta área, por lo tanto, mantenerlo estable depende de la ejecución de muchas acciones, incluyendo la de gestión con elemento tercero, como por ejemplo si algún servicio se brinda contratando un segmento de redes de otro operador, debemos de velar por la estabilidad de esta.

¿Entiende que es necesario la implementación de algún sistema que le ayude a mejorar sus procesos internos, cuál/es proceso/s?

Hay muchos campos que recorrer para mejorar el desempeño de la gestión del área, el desempeño de nuestra área depende de acción que son reactiva y proactivas, estos término son relativos, reaccionar al tiempo o cuando el problema está pequeño y todavía no te afecta, es en sí una acción preventiva, preventiva del problema mayor, estoy corrigiendo un problema menos reactivamente pero al mismo tiempo estoy actuando preventivamente para que no se produzca el problema mayor, a medida que yo pueda incorpora herramienta de predicción y pueda atacar en su etapa temprano, estaré elevando mi nivel de servicio y eso solo se puede lograr con herramientas, herramienta que tenga la capacidad de correlacionar eventos y encontrar patrones o predecir que algo va a pasar con altas o bajas probabilidades.

¿Poseen alguna plataforma que se encargue de monitorear el estado de salud de los elementos(hardware) de la red móvil de WIND Telecom?

Si, existe tanto de la red móvil como para todos los elementos de la red.

¿Esa plataforma es propietaria de WIND Telecom o provista por un 3ero?

Las mayorías son provistas por un tercero, pero tiene pequeñas herramientas periférica que hemos desarrollado sobre ella para optimizar lo que ellas hacen.

¿Podemos tener acceso a los datos que provee esa plataforma?

Si

¿Poseen alguna fuente de datos donde podamos ver la disponibilidad de los nodos de la red móvil de WIND Telecom, detallando las causas? ¿Podemos tener acceso a esos datos y que tanto histórico poseen?

- Si, existe un sistema para colectar esas informaciones
- Sin duda pueden tener acceso a estas informaciones
- Hay informaciones que tiene más de tres años disponibles. Casi todas las informaciones de niveles de disponibilidad y desempeño de la red móvil tienen varios años en que se han ido colectando

¿Poseen alguna fuente de datos donde se detalle el costo operativo que con lleva solucionar las averías de la red móvil? ¿Podemos tener acceso a esos datos y que tanto histórico poseen?

Existe archivos planos, archivos que incorpora información de diferente fuente (finanza y otros sistemas), para la red móvil existe suficientes datos de los costos de operar, de mantenerla y los costó de actualizarla, estos datos se tienen con los fines de entender la rentabilidad de los servicios que se monta en esas redes y la rentabilidad de los nodos cuando se construyen. Existe también hoja de cálculo en excel que contienen información detallada de cada nodo, especialmente lo de la red móvil. Si se puede tener acceso a la misma.

¿Poseen alguna fuente de datos que sirva para analizar la satisfacción del cliente con el servicio de WIND Telecom, podemos tener acceso a esos datos, que tanto histórico poseen?

Si, tenemos dos tipos de datos, primero indicadores internos (mediciones internas) que fruto de un análisis exhaustivo se ha identificado que impacta la satisfacción de nuestro cliente que son datos obtenidos por variables de reclamación o nodo de red y el otro tipo de datos viene por encuestas (mediciones externas), la cual nosotros correlacionados.

Estos datos de mediciones internas los producen operaciones de servicios, pero se alimenta en parte de datos que proveemos nosotros del área de tecnología que son reclamaciones, datos de uso de la red, datos de disponibilidad la cuales son recorrido y relaciona con datos de producto, competencia, niveles de servicio fuera de la red (servicio directo oficina, call center o punto de venta). Por otro lado, las mediciones externas se realizan por instituciones tercero o nuestros call center bajo un estándar.

Hay un histórico de tres años de informaciones, se puede obtener a través del área de tecnología o el área de servicio.

¿Poseen alguna fuente de datos donde podamos ver la tasa de deserción (churn) de los clientes, clasificada por causas, podemos tener acceso a esos datos, que tanto histórico poseen?

Si, ese es un dato que se mide periódica e ininterrumpidamente por el periodo de 5 años y con suficientes niveles de desglose, observando dos tipos de deserción la que es por voluntad del cliente y la realizada por la empresa, usualmente por incumplimiento del cliente. El dato es producido por operaciones de servicios y se puede obtener tanto por ellos como por el departamento de tecnología.

¿Cree que sería útil para la empresa o su área una herramienta para predecir posibles averías en su red móvil, ¿Qué beneficios entiende le aportaría dicha herramienta?

El primero es elevar el nivel de servicio y satisfacción de nuestro cliente y la segunda es optimizar nuestra operación y este se fundamenta en que está claramente determinado que es más barato prevenir que corregir, independiente mente que nosotros mejoramos la satisfacción también debería reducir costo de operaciones.

¿Tiene conocimiento del proceso macro para la adquisición de elemento?

Hay dos elementos, el elemento de la red (infraestructura de la red) y los elemento que se utiliza punto fundamental para la entrega de servicio que se le vende al cliente, que es el dispositivo sobre el cual el cliente maneja dichos servicios. en ambo caso existe un proceso similar al inicio y es que en nuestra área de ingeniería y planificación de tecnología determina y homologa los elementos que se van a adquirir basado el lineamiento estratégico que se han determinado previamente, esto fruto de una evaluación exhaustiva de diferente fabricante, una vez seleccionado la siguiente compra no se somete a una evaluación de nuevo fabricante, por lo tanto, cada elemento ya tiene predefinido el fabricante. ya para la adquisición lo maneja el departamento de compra con las especificaciones de ingeniería tanto en el elemento como en el volumen.

En el caso de los dispositivos, compra maneja el volumen a diferencia de los elementos de la red.

¿Cree que exista otro sistema dentro de la empresa donde se podría implementar la predicción de fallas, podría mencionar alguno?

Sin ninguna duda, tenemos red móvil para ofrecer servicios a los clientes masivos (residencial e individuales) pero hay toda una serie de elemento de red para proveer servicio a

clientes de mayor valor y eso elementos de red tienen el reto para nosotros de no posees sistema de gestión para hacer predicciones, el reto es tomar datos que están mucho más crudo y convertirlos en informaciones relevante para tomar acciones. Cualquier acción de predicción que se vaya a hacer para la red móvil tiene aplicación sobre los otros elementos de la red, incluyendo el elemento menos inteligente como por ejemplo lo elemento de potencia o climatización.

2. Resultado entrevista a Elvy Durán

¿Nombre? Elvy Durán
¿Área de labor? Operaciones de la red
¿Tiempo laborando en la empresa? 10 años
¿Cargo desempeña? Director de operaciones de la red
¿Funciones que desempeña?

Todo el esquema de operación, mantenimiento general de la red, la predicción del funcionamiento de la red y todo el aspecto correctivo y preventivo.

¿Qué tipo de informaciones manejan en su área?

De la inmediata: todos los eventos de alarmas, trouble tickets, todos los temas estadísticos de la red, performance de la red, todos los temas de plataforma de TI y también algunos temas relacionados a los impactos en ventas y desarrollo de productos.

¿Estas informaciones la produce el área o la consumen de otras áreas o sistemas?

La mayoría la produce el área, una parte viene del área de mercadeo (un porciento muy bajo) pero al final toda la plataforma que las producen son responsabilidad nuestra.

¿Cómo se almacenan las informaciones producida por el área? (físico o digital) (Base de datos o archivos)? Se almacenan en base de datos.

¿Las informaciones de otras áreas o sistemas provienen de Base de datos o archivos?

Se copilan prácticamente todas en base de datos, la más reciente llega como información plana, básicamente la parte de las redes sociales, que se está implementando recientemente, pero

al final es convertida en un esquema de base de datos. De las informaciones que utilizan otras áreas y su área en sí, se consumen dando resultados y tendencias para análisis, se manejan en combinación, entre base de datos o archivos locales pero la mayoría terminan manejándose en base de datos, si hay muchas todavía que están moviéndose en las dos aguas, una parte aquí y otra parte en base de datos, una parte en archivo puro de excel, algunos archivos planos pero el resultado final termina en una base de datos.

¿Utilizan alguna herramienta para la realización de sus funciones diarias, cual/es?

Parte de las herramientas básicas de la informática están los relacionados con temas de reporting, ya los temas de administración de base de datos y los temas de power BI.

¿Esas herramientas producen las informaciones consumidas por el área o usan como fuentes esas informaciones producidas por el área?

Depende la información, pero sí, se alimentan de diferentes fuentes. Otras internas producidas por el área nuestra y otras producidas por el área de mercadeo.

¿Sabe usted cómo su área impacta en la calidad del servicio ofertado a los clientes?

Sí. En un grado de importancia del 1 al 10 le doy un 11.

¿Sabe usted cómo su área impacta en el proceso de solución de las averías en la red móvil?
Sí.

¿Entiende que es necesario la implementación de algún sistema que le ayude a mejorar sus procesos internos, cuál/es proceso/s?

De cara a la calidad y sobre todo a la solución de averías, si debiéramos tener algunos elementos adicionales que nos permitieran sobre todo manejar, quizás con más facilidad, todos los temas de eventos en la red, los tickets manejados sobre todo y ahora mismo el tema de poder correlacionar con más facilidad. Esos temas: los tickets, los eventos en la red y los reportes de clientes.

¿Poseen alguna plataforma que se encargue de monitorear el estado de salud de los elementos(hardware) de la red móvil de WIND Telecom?

Sí, la overall de la red, el Network Management System (NMS) persé de la plataforma de la red. Si nos vamos a los hierros de la red móvil como tal, ahora el servicio de la red que damos tiene otros elementos detrás que dependen, influyen significativamente en el servicio de la red móvil, ahora sí, la red móvil persé, intrínseca ella, ese CORE como las BTS, tienen ellas un NMS exclusivo para ellas. El resto de la red descansa también sobre otra plataforma sombrilla que maneja todos los eventos que se manejan en ella y lo mismo pasa con las herramientas, las partes horizontales que sustentan el servicio en las de TI.

¿Esa plataforma es propietaria de WIND Telecom o provista por un 3ero?

Las de monitoreo todas son provistas por terceros. Son propiedad de WIND ya, WIND las adquirió, pero son provistas por terceros, todas las que son con el monitoreo de la red y manejos de trouble ticket de la red. La parte de CRM y de manejo de trouble ticket de clientes finales es un desarrollo in house.

¿Podemos tener acceso a los datos que provee esa plataforma?

Sí, entiendo que sí, porque todas esas plataformas manejan base de datos y la mayoría, no todas, manejan escenarios de web service o manejan escenario de SAP que pueden permitir extraer las informaciones sin muchas complicaciones.

¿El acceso que nos darían a esa información, tendrían condicionantes como es de esperarse?

Claro, hay que firmar un acuerdo de confidencialidad obligatorio.

¿Poseen alguna fuente de datos donde podamos ver la disponibilidad de los nodos de la red móvil de WIND Telecom, detallando las causas? ¿Podemos tener acceso a esos datos y que tanto histórico poseen?

Sí, existe. Todavía un poco manual el proceso, pero existe una base de datos donde se tiene parte de esas informaciones, de hecho, sería parte del proceso que aún necesitamos automatizar más para eficientizar los procesos de correlación.

El histórico va a depender de la plataforma, de disponibilidad se posee el histórico full de la empresa. Si nos vamos a la parte de disponibilidad de red móvil puntualmente, desde que esta red se instaló en el 2015.

¿Poseen alguna fuente de datos donde se detalle el costo operativo que con lleva solucionar las averías de la red móvil? ¿Podemos tener acceso a esos datos y que tanto histórico poseen?

No, detallado no se tiene. Lo que tenemos son los costos generales por departamento y tú pudieses tratar de inferir los. El tema aquí es que no existe una correlación entre los casos puntuales de la red móvil y el costo.

Un ejemplo muy sencillo: tú tienes un departamento de operaciones y mantenimiento de la red, o quizás el mismo centro de operaciones que es más complejo todavía pero un departamento de operaciones de la red, que ejecuta un mantenimiento preventivo y ejecuta un mantenimiento correctivo, ahora qué sucede, cuando el empleado va a hacer el mantenimiento preventivo o va a ejecutar un proceso correctivo, ese grupo no solo trabaja parte de la red móvil, ese grupo tiene otros roles. Un ejemplo es que maneja partes de sus costos están relacionados con la red de transporte IP que no es persé red móvil, esa red de transporte IP no sólo maneja la red móvil, maneja la red de servicios de negocios, maneja otros tipos de servicios como el de voz, entonces cuando el empleado va y trabaja una avería o una implementación de la red móvil o un servicio de un cliente de negocios, ese proceso, detrás en la parte contable, no está separado. Yo no tengo forma de decir que estas dos horas que ese empleado trabajó, las trabajó en la red móvil o no, no hay ese mecanismo. Lo que, si pudiese hacerse, pero hay que calcularlo, y se calcula cada cierto tiempo, sobre todo cuando hay que hacer temas presupuestarios es, que tu tratas de hacer más o menos una correlación, entre el volumen de horas que pudo dedicarles un técnico a los diferentes procesos de mantenimiento correctivo y preventivo versus los costos generales del departamento. Es como si tu hicieras un ponderado, tú dices bueno, este departamento por los tickets de avería o por el programa de mantenimiento preventivo, tú dices cuántos tickets trabajó en mantenimientos móviles, cuántos trabajó en clientes de negocios, cuántos trabajó en servicios voz, etc, y cada tickets de esos tiene un tiempo trabajado, entonces tú dices ok, pondera las horas, para decirlo así, y pudiese asumir un ponderado del costo que ese departamento está influyendo en la red móvil.

El problema es que el dato como tal no existe, sino que hay que crearlo, se saca de manera muy puntual. Sería un plus en posible mejora.

¿Poseen alguna fuente de datos que sirva para analizar la satisfacción del cliente (quality of experience) con el servicio de WIND telecom?

Quality of experience no existe en la empresa, no existe una plataforma. Quality of experience es una plataforma automática que simula un cliente, pero si existen las encuestas de satisfacción. Se hacen cada vez que un cliente llama al call center se cierra con una pequeña encuesta de satisfacción y dos veces al año, si no me falla la memoria, también se hace una encuesta de satisfacción al cliente. Se contrata una empresa externa que hace un muestreo y da ese resultado y ese resultado existe en una serie de documentaciones, no está tabulado, o sea no está en base de datos, mejor dicho.

El resultado de las encuestas de satisfacción de cliente que se maneja dos veces al año las maneja mercadeo analítico. El resultado de la encuesta on time en la llamada del cliente la maneja José Raúl.

¿Poseen alguna fuente de datos donde podamos ver la tasa de deserción o churn rate de los clientes de WIND telecom? ¿Podemos tener acceso a esos datos y que tanto histórico poseen?

Como base de datos no existe como tal, pero se manejan archivos externos. Se le da un seguimiento diario, pero se maneja en archivos externos.

Sí, se poseen históricos y fácilmente hay datos por ahí hasta del 2010.

¿Cree que sería útil para la empresa o su área una herramienta para predecir posibles averías en su red móvil, ¿Qué beneficios entiende le aportaría dicha herramienta?

Sí, es interesante. Depende el alcance que le queramos dar y cuál es el scope general que se le quiera poner pero sería interesante porque al tu poder tratar de predecir o si tu logras predecir

con antelación, para decir un dado tonto, con tu predecir el 30 o 40 por cientos de los eventos que tu puedas tener en tu red, solamente viendo los eventos de impacto al servicio como tal, ya eso te debe reducir significativamente tus reportes de quejas de los clientes.

Solamente por ahí debe pagar solito, pero ahora súmale que si tu logras predecir con suficiente tiempo una salida se pudiera reducir en un 30% el costo operativo.

Pongamos el escenario en que logras predecir una salida, pero si el tiempo con el que logras predecir no es significativo pues no gano nada con predecir, predecir un evento que se va a salir mañana a las 12 de la noche de hoy nos genera el mismo impacto que si no lo predijese, ahora si yo puedo predecirlo con suficiente tiempo de antelación es muy probable o seguro que primero, el tiempo de ir a corregirlo me da:

a. Más holgura
b. No tengo que trabajarlo en horas no laborables lo que debería reducir mi costo de horas extras y adicionales
c. El impacto más grande en donde pescaría más y sin miedo, sería un impacto significativo en la percepción del cliente. El cliente se va a sentir más satisfecho, vas a tener menos impactos en los servicios y por ende esto se traduce en un cliente con más fidelidad. Y eso simple y llanamente al final se traduce en lo más importante para cualquier empresa, que es el crecimiento de su cartera y por ende todo eso es finanzas. Dinero entra y poco dinero sale. Cada cliente que hace churn en cualquier empresa es un cliente del cual yo perdí dinero, no solamente porque dejé de percibir un ingreso, sino que dependiendo el tiempo en que el haga el churn, estoy perdiendo capital de inversión. ¿En qué sentido? Pues cuando yo a un cliente le doy un equipo, que se lo regalé para que entrara a la empresa, ese equipo tiene un costo, y tiene un tiempo para yo recuperarlo en el concepto de la renta, yo digo, esta renta así que me gano por ejemplo $20.00 pesos de cara a mis costos generales, de esos $20.00 pesos hay $5.00 que yo los voy a usar para cubrir mi equipo, así es que lo reflejo de manera contable, ahora bien si ese cálculo me da que ese cliente tiene que durar 1 año conmigo o 2 años y si se ve en 3 o 4 meses, el resto de los meses es un dinero perdido, tomando la condición de que hay muchos clientes que

retiran el servicio y no me devuelven el equipo y se lo quedan. Pero igual, si les doy un equipo nuevo, aunque me lo devuelvan, ese equipo cuando lo vuelva a poner a disposición de ventas ya no va al mismo precio ya que es un equipo que en el mejor escenario se lo puedo vender como refurbished aunque nunca le haya hecho reparación y pierde un valor significativo el equipo.

¿Cree que exista otro sistema dentro de la empresa donde se podría implementar la predicción de fallas, podría mencionar alguno?

Dada la naturaleza de la empresa, el hecho de tu ir a predecir las fallas en la red móvil, insisto, va a depender del scope, porque si yo lo que quiero predecir es nada más la falla del hierro como tal, me quedé sólo en la red móvil DTS con de la red. Ahora, si yo lo que quiero es predecir la falla del servicio y no del hierro tengo que obligatoriamente que incluir en el proceso de predicción otras plataformas como, por ejemplo:

- La red de transporte IP
- La que mucha gente no escucha, pero es crítica, la energía eléctrica
- Que la gente la oye menos y cree que no existe porque a veces es intangible, toda la parte de tecnología de la información que sostiene en los servicios, en nuestro caso eso va desde firework hasta los propios servidores que manejan las aplicaciones y almacenan la misma data que se debe utilizar para poder predecir las cosas. La misma plataforma que monitorea los eventos, si no están estables, si no tienen un buen performance, tampoco voy a poder tener la información, ni visual del técnico del centro de operaciones de la red ni en cualquier plataforma de predicción que yo quiera hacer para poder hacer un análisis real del tema, en el caso WIND telecom, yo entiendo, que si se hace en el escenario con el que debe hacerse, tendría que abarcar varias cosas de la empresa obligatoriamente, aunque se puede arrancar en fases, obligatoriamente, completa es difícil ya que estamos hablando de diferentes sistemas y habría que ver la disponibilidad de datos de esos diferentes sistemas. Insisto con el scope, yo puedo decir, y a esto le voy a llamar la predicción de la disponibilidad, para decir algo, yo quiero decir qué tanto falla una BTS, pero desde el punto de vista del hierro la BTS falla y se le daña a ella una pieza, la BTS

falla sólo si se apaga, pero desde el punto de vista del servicio la BTS está fuera, no sirve si la red de transporte no sirve, la que tu menos, porque era menos tangible vas a meter es la parte de informática, la parte de TI, servidores, etc., porque eso es más horizontal en todo el escenario y su impacto regularmente es catastrófico. Es igual que el power puntual, el power un generador que falle con un sistema de backup ineficiente, es un colapso. Pero, de hecho, hay que comenzar de arriba hacia abajo, y en la pirámide, la parte de arriba es la red móvil, pero igual, estamos hablando de la red móvil pero en el mismo nivel de la red móvil como pirámide, funcionalidad y de servicio tú tienes la plataforma de fibra óptica GPON, viendo servicios masivos, pero concentrémonos en LTE, la parte móvil, la punta de la pirámide sí, son los hierros de la BTS, lo que la gente ve en la calle como una antena, eso ahí, pero si tú quieres hacer una predicción real, hay que ver eso y una serie de componentes adicionales, que es lo que permite que eso funcione.

3. Resultado entrevista a Alan Molina

¿Nombre? Alan Molina

¿Área de labor? Ingeniería de la Red

¿Tiempo laborando en la empresa? 7 años

¿Cargo desempeña? Gerente de Ingeniería y Optimización de la Red LTE

¿Funciones que desempeña?

a. Instalación y puesto en funcionamiento de las nuevas localidades
b. Supervisión del core de equipos LTE
c. Supervisión de los equipos móviles entregados a los clientes
d. Soporte a los departamentos de mantenimiento y operaciones en la red LTE
e. Optimización de la red LTE

¿Qué tipo de informaciones manejan en su área?

Parámetros de los equipos, configuraciones y los kpi

¿Estas informaciones la produce el área o la consumen de otras áreas o sistemas?

Estas informaciones se producen tanto en su área como en otras áreas

¿Cómo se almacenan las informaciones producida por el área? (físico o digital) (Base de datos o archivos)?

Esta información se maneja en excel

¿Las informaciones de otras áreas o sistemas provienen de Base de datos o archivos?

Provienen de base de dato, llega en excel.

¿Utilizan alguna herramienta para la realización de sus funciones diaria, cual/es?

PowerBI y Excel para análisis de la data y en la parte de optimización herramienta para los drives test y post procesamiento.

¿Esas herramientas producen las informaciones consumidas por el área o usan como fuentes esas informaciones producidas por el área?

En ingeniería y optimización producen los datos. De cara a los análisis y la data que consumen es datos que provienen de otro departamento.

¿Sabe usted cómo su área impacta en la calidad del servicio ofertado a los clientes?

En la medida que los recursos se están agotando el departamento de Ingeniería de la Red debe ir haciendo ajustes para que el cliente no se vea afectado en su servicio

¿Sabe usted cómo su área impacta en el proceso de solución de las averías en la red móvil?

Son el soporte segundo nivel en las celdas y atienden las averías a nivel de core. Se encargan de todo lo relacionado a Radio Frecuencia.

¿Entiende que es necesario la implementación de algún sistema que le ayude a mejorar sus procesos internos, cuál/es proceso/s?

Entiende que sí. Entre los procesos que deben mejorar son los KPI, la información de Ingeniería que manejen y propone una implementación de una BD y alguna herramienta para

analizar la data para facilitar la información. En la parte de optimización necesitan una herramienta para KPI, pero no tienen los recursos y tiempos para dar respuesta en el tiempo requerido.

¿Poseen alguna plataforma que se encargue de monitorear el estado de salud de los elementos(hardware) de la red móvil de WIND Telecom?

Si, utilizan Netnumen RAN y Netnumen CORE para la parte del NOC de ahí sale la data cruda para los KPI y también se definen KPI.

¿Esa plataforma es propietaria de WIND Telecom o provista por un 3ero?

Esta plataforma es propietaria de WIND desarrollada por el suplidor de la infraestructura (ZTE).

¿Podemos tener acceso a los datos que provee esa plataforma?

Si, con la firma de contrato de confidencialidad.

¿Poseen alguna fuente de datos donde podamos ver la disponibilidad de los nodos de la red móvil de WIND Telecom, detallando las causas? ¿Podemos tener acceso a esos datos y que tanto histórico poseen?

Del CORE se guardan hasta 3 meses con la data cruda, no cree que haya más información superior a los 6 meses. La información está cruda y se puede hacer resúmenes.

¿Cree que sería útil para la empresa o su área una herramienta para predecir posibles averías en su red móvil, ¿Qué beneficios entiende le aportaría dicha herramienta?

Indirectamente si, directamente no. El sistema de RF por un tema de interferencia para poder decir que una celda pueda tener algún tipo de inconveniente o KPI que pudiera facilitar su labor. El beneficio sería mantenimiento preventivo y al final de cuenta se transforma en calidad de servicio al usuario.

¿Cree que exista otro sistema dentro de la empresa donde se podría implementar la predicción de fallas, podría mencionar alguno?

Si, en desarrollo y de tecnología de la información. Además de GPON y Electromecánica.

4. Resultado entrevista Bin Hon NG

¿Nombre? Bing NG

¿Área de labor? Centro de Operaciones de la Red (NOC)

¿Tiempo laborando en la empresa? 11 años

¿Cargo desempeña? Gerente del Centro de Operaciones de la Red

¿Funciones que desempeña? Dirigir el NOC

¿Qué tipo de informaciones manejan en su área?

Averías de la empresa, disponibilidad de los servicios, operabilidad de los equipos de la red, entrada de las averías en la red móvil y voz.

¿Estas informaciones la produce el área o la consumen de otras áreas o sistemas?

Es una combinación. Las producidas por el NOC son: entrada de averías, visibilidad de la red, las recepciones de las alertas o alarmas. La estadística de disponibilidad y performance se obtienen de otro departamento.

¿Cómo se almacenan las informaciones producida por el área? (físico o digital) (Base de datos o archivos)?

Todo se maneja digital una parte en BD o otra en Excel.

¿Las informaciones de otras áreas o sistemas provienen de Base de datos o archivos?

Se maneja mixto en BD y Excel.

¿Utilizan alguna herramienta para la realización de sus funciones diara, cual/es?

Monitoreo con herramientas básicas propias de ZTE (Netnumen RAN y CORE) y las demás son para el manejo de las informaciones con plantillas de excel. Para el monitoreo de tráfico se utiliza PRTG. No utilizan PowerBI, pero están interesados.

¿Sabe usted cómo su área impacta en la calidad del servicio ofertado a los clientes?

Si en cuanto a la red móvil y de transporte su departamento juega un papel crucial en la solución de las averías.

¿Sabe usted cómo su área impacta en el proceso de solución de las averías en la red móvil?

Son los encargados de despachar el recurso para solucionar las averías en la red.

¿Entiende que es necesario la implementación de algún sistema que le ayude a mejorar sus procesos internos, cuál/es proceso/s?

Poseen una debilidad para visualizar la estadística de tiempo de solución, causa más frecuente, así como tener a mano la disponibilidad de la red. Poder ver los históricos le ayudarían bastante para resolver las fallas de la red.

¿Poseen alguna plataforma que se encargue de monitorear el estado de salud de los elementos(hardware) de la red móvil de WIND Telecom?

Si, todo se realiza con la herramienta de ZTE.

¿Esa plataforma es propietaria de WIND Telecom o provista por un 3ero?

Propietaria de WIND Telecom, pero provista por ZTE.

¿Podemos tener acceso a los datos que provee esa plataforma?

Se puede ver y extraer dicha data sin problemas.

¿Poseen alguna fuente de datos donde podamos ver la disponibilidad de los nodos de la red móvil de WIND Telecom, detallando las causas? ¿Podemos tener acceso a esos datos y que tanto histórico poseen?

Actualmente usan HP Service Manager para los reportes de avería y es con reportes salida de ahí conocen la cantidad de avería. Si es para conocer puntos puntuales de los nodos, con la herramienta de ZTE se obtiene datos precisos. El histórico es de solo seis meses, hay que validar en la base datos de la plataforma para ver si hay forma de acceder más tiempo. En service Manager se obtienen 7-8 años.

¿Cree que sería útil para la empresa o su área una herramienta para predecir posibles averías en su red móvil, ¿Qué beneficios entiende le aportaría dicha herramienta?

Si, le ayudara bastante para la solución de los casos a manera de prevención. Más le compete al departamento en los campos, para mejorar la gestión de los recursos.

¿Cree que exista otro sistema dentro de la empresa donde se podría implementar la predicción de fallas, podría mencionar alguno?

En el área de mantenimiento y en ingeniería y optimización de la red móvil.

5. Resultado entrevista a Luisa Perez

¿Nombre? Luisa Perez
¿Área de labor? Compra y Registro
¿Tiempo laborando en la empresa? 9 años
¿Cargo desempeña? Gerente senior de compras y logística
¿Funciones que desempeña?

Dirigir el proceso de abastecimiento (solicitud de compras, compra, almacenamiento y transporte)

¿Qué tipo de informaciones manejan en su área?

Manejan los requerimientos de los clientes. Hay elementos que se manejan en inventario y existen otros que no se colocan en inventario, por ejemplo, elementos usados para proyectos de expansión de la red entre otros.

¿Estas informaciones la produce el área o la consumen de otras áreas o sistemas?

Producen estadísticas, pero sus operaciones dependen de información de otras áreas.

¿Cómo se almacenan las informaciones producida por el área? (físico o digital) (Base de datos o archivos)?

Están en archivo de Excel, algunas históricas están en un drive.

¿Utilizan alguna herramienta para la realización de sus funciones diaria, cual/es?

SAP y Ofimática.

¿Esas herramientas producen las informaciones consumidas por el área o usan como fuentes esas informaciones producidas por el área?

Produce información, ya que todo se debe entrar de forma manual pues hay que digitarla. El proceso inicia con un pedido y todo debe ser digitado.

¿Sabe usted cómo su área impacta en la calidad del servicio ofertado a los clientes?

Es el departamento que se encarga de entregar los insumos para entregar los servicios que los clientes requieren con la calidad esperada.

¿Sabe usted cómo su área impacta en el proceso de solución de las averías en la red móvil?

Si, compras se encarga de todo el proceso de gestión de los elementos que se usan para la solución de las averías.

¿Entiende que es necesario la implementación de algún sistema que le ayude a mejorar sus procesos internos, cuál/es proceso/s?

Facilita bastante poder ampliar la base de usuario que puedan tener para usar SAP.

¿Nos puede detallar brevemente el proceso de adquisición de elementos(hardware) para la red móvil de WIND Telecom?

1. Área define lo que desea (Mayormente ingeniería búsqueda en el mercado y homologación del producto)
2. Debe llegar con la matriz de aprobación
3. Deben buscar los productos que cumplan con los montos topes aprobados para cumplir con auditoría
4. Se procede a realizar una orden de compra
5. Se regresa al área que solicita para validar que la orden de compra sea lo que se requiere y luego se aprueba
6. Luego pasa nuevamente por la matriz de aprobación

7. Se coloca en SAP
8. Se valida por vendedor y por encargado de Compras
9. Se solicita
10. Se provisiona

Para averías:

a. No se piden muchas para inventario de avería
b. Se hizo un análisis hace un tiempo de cuellos de botella, pero a la fecha la empresa no posee inventario de equipos que son de alto impacto en la red.

¿Este proceso implica algún costo para la empresa?

No posee costo para la empresa. Representan el 3-4% del costo total en la empresa. En vista de almacén si posee, pues flete de avión pudiera hasta aumentar un 30%. Y este costo se coloca a la avería(gasto) o al proyecto(inversión).

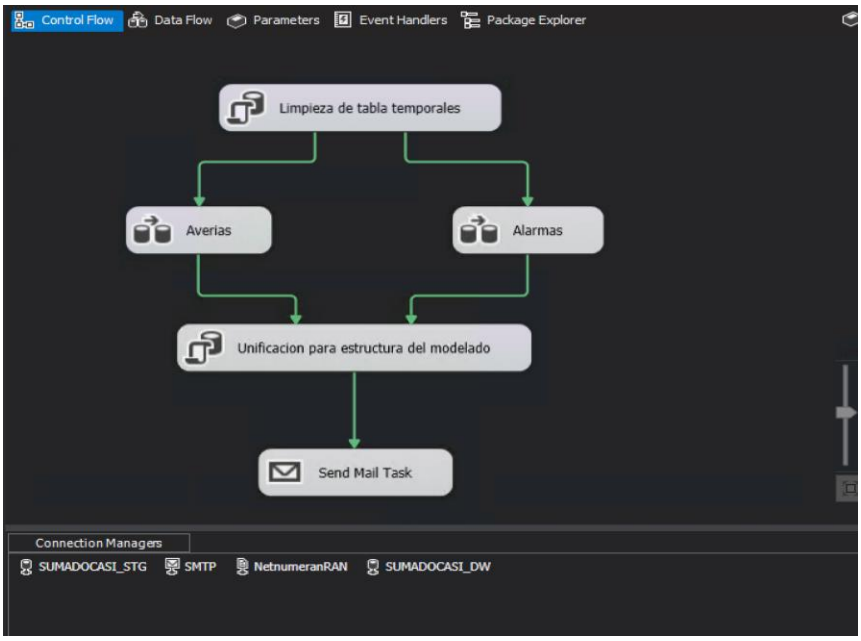
¿Poseen alguna fuente de datos donde se detalle los costos asociados a este proceso?

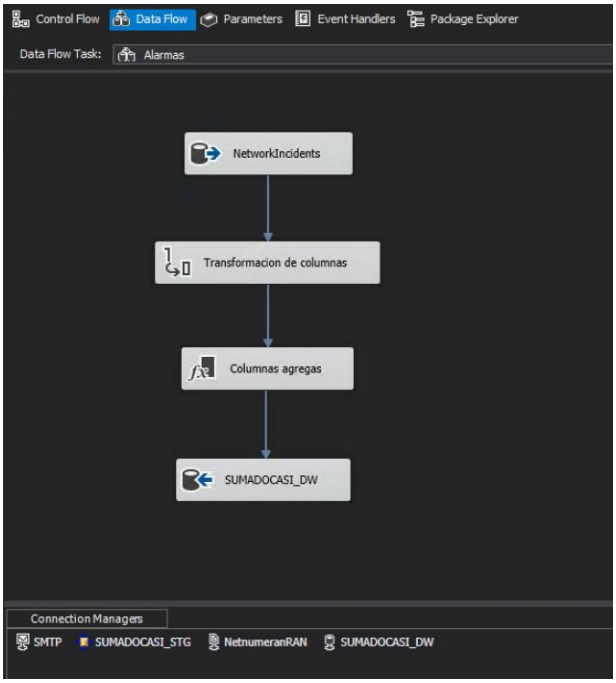
¿Podemos tener acceso a esos datos, que tanto histórico poseen?

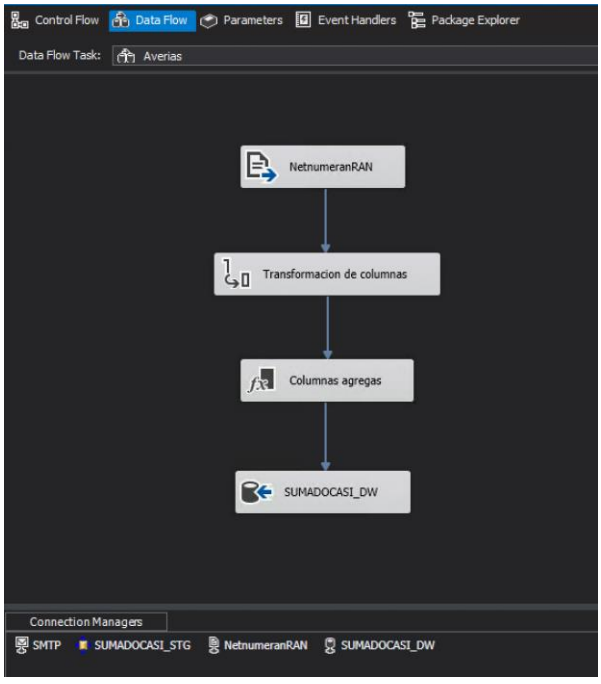
No poseen

Anexo B - Resultado de Modelos predictivos

Procesos ETL (SSIS) para transformación y carga del dataset para el modelo







Análisis Exploratorio de Datos (EDA)

Overview

Dataset info

Number of variables	37
Number of observations	3013
Total Missing (%)	0.0%
Total size in memory	871.0 KiB
Average record size in memory	296.0 B

Variables types

Numeric	30
Categorical	2
Boolean	2
Date	0
Text (Unique)	0
Rejected	3
Unsupported	0

Variables

198092014
Numeric

Distinct count	4	Mean	0.012944
Unique (%)	0.1%	Minimum	0
Missing (%)	0.0%	Maximum	3
Missing (n)	0	Zeros (%)	99.1%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

Statistics | Histogram | Common Values | Extreme Values

Quantile statistics

Minimum	0
5-th percentile	0
Q1	0
Median	0
Q3	0
95-th percentile	0
Maximum	3
Range	3
Interquartile range	0

Descriptive statistics

Standard deviation	0.1508
Coef of variation	11.65
Kurtosis	244.75
Mean	0.012944
MAD	0.025647
Skewness	14.564
Sum	39
Variance	0.022741
Memory size	23.6 KiB

198092029
Numeric

Distinct count	5	Mean	0.025888
Unique (%)	0.2%	Minimum	0
Missing (%)	0.0%	Maximum	4
Missing (n)	0	Zeros (%)	98.7%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092044
Numeric

Distinct count	21	Mean	0.28244
Unique (%)	0.7%	Minimum	0
Missing (%)	0.0%	Maximum	98
Missing (n)	0	Zeros (%)	93.4%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092053
Numeric

Distinct count	39	Mean	0.74411
Unique (%)	1.3%	Minimum	0
Missing (%)	0.0%	Maximum	385
Missing (n)	0	Zeros (%)	96.7%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092074
Numeric

Distinct count	24	Mean	0.30667
Unique (%)	0.8%	Minimum	0
Missing (%)	0.0%	Maximum	66
Missing (n)	0	Zeros (%)	93.9%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092091
Numeric

Distinct count	12	Mean	0.059409
Unique (%)	0.4%	Minimum	0
Missing (%)	0.0%	Maximum	30
Missing (n)	0	Zeros (%)	99.4%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092096
Numeric

Distinct count	10	Mean	0.027215
Unique (%)	0.3%	Minimum	0
Missing (%)	0.0%	Maximum	21
Missing (n)	0	Zeros (%)	99.3%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092097
Numeric

Distinct count	34	Mean	0.35314
Unique (%)	1.1%	Minimum	0
Missing (%)	0.0%	Maximum	51
Missing (n)	0	Zeros (%)	97.3%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

~~198092099~~
Highly correlated

This variable is highly correlated with `198092096` and should be ignored for analysis

Correlation 0.92242

198092108
Numeric

Distinct count	14	Mean	0.12811
Unique (%)	0.5%	Minimum	0
Missing (%)	0.0%	Maximum	17
Missing (n)	0	Zeros (%)	97.1%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

~~198092117~~
Highly correlated

This variable is highly correlated with `198092074` and should be ignored for analysis

Correlation 0.94485

198092295
Numeric

Distinct count	71	Mean	2.3375
Unique (%)	2.4%	Minimum	0
Missing (%)	0.0%	Maximum	330
Missing (n)	0	Zeros (%)	86.4%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092348
Numeric

Distinct count	10	Mean	0.080651
Unique (%)	0.3%	Minimum	0
Missing (%)	0.0%	Maximum	19
Missing (n)	0	Zeros (%)	95.9%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092458
Boolean

Distinct count	2	Mean	0.0023233
Unique (%)	0.1%		
Missing (%)	0.0%		
Missing (n)	0		

Toggle details

198092550
Numeric

Distinct count	76	Mean	2.2961
Unique (%)	2.5%	Minimum	0
Missing (%)	0.0%	Maximum	131
Missing (n)	0	Zeros (%)	90.0%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092551
Numeric

Distinct count	17	Mean	0.25357
Unique (%)	0.6%	Minimum	0
Missing (%)	0.0%	Maximum	18
Missing (n)	0	Zeros (%)	93.7%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092553
Numeric

Distinct count	62	Mean	1.5795
Unique (%)	2.1%	Minimum	0
Missing (%)	0.0%	Maximum	123
Missing (n)	0	Zeros (%)	93.9%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092554
Numeric

Distinct count	16	Mean	0.10189
Unique (%)	0.5%	Minimum	0
Missing (%)	0.0%	Maximum	28
Missing (n)	0	Zeros (%)	98.9%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198092556
Numeric

Distinct count	42	Mean	2.0899
Unique (%)	1.4%	Minimum	0
Missing (%)	0.0%	Maximum	1408
Missing (n)	0	Zeros (%)	97.4%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198094830
Numeric

Distinct count	22	Mean	0.29406
Unique (%)	0.7%	Minimum	0
Missing (%)	0.0%	Maximum	37
Missing (n)	0	Zeros (%)	90.1%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198094832
Numeric

Distinct count	96	Mean	5.0604
Unique (%)	3.2%	Minimum	0
Missing (%)	0.0%	Maximum	654
Missing (n)	0	Zeros (%)	82.6%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198094833
Numeric

Distinct count	15	Mean	0.16893
Unique (%)	0.5%	Minimum	0
Missing (%)	0.0%	Maximum	22
Missing (n)	0	Zeros (%)	93.2%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198094856
Numeric

Distinct count	3	Mean	0.011616
Unique (%)	0.1%	Minimum	0
Missing (%)	0.0%	Maximum	32
Missing (n)	0	Zeros (%)	99.9%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198094858
Numeric

Distinct count	16	Mean	0.3299
Unique (%)	0.5%	Minimum	0
Missing (%)	0.0%	Maximum	45
Missing (n)	0	Zeros (%)	96.2%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198094863
Numeric

Distinct count	23	Mean	0.30733
Unique (%)	0.8%	Minimum	0
Missing (%)	0.0%	Maximum	37
Missing (n)	0	Zeros (%)	92.3%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198096551
Numeric

Distinct count	4	Mean	0.018918
Unique (%)	0.1%	Minimum	0
Missing (%)	0.0%	Maximum	3
Missing (n)	0	Zeros (%)	98.6%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198097050
Numeric

Distinct count	32	Mean	0.49187
Unique (%)	1.1%	Minimum	0
Missing (%)	0.0%	Maximum	108
Missing (n)	0	Zeros (%)	87.6%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

198097172
Boolean

Distinct count	2	Mean	0.0003319
Unique (%)	0.1%		
Missing (%)	0.0%		
Missing (n)	0		

Toggle details

198098465
Numeric

Distinct count	46
Unique (%)	1.5%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0

Mean	1.0216
Minimum	0
Maximum	71
Zeros (%)	89.3%

Toggle details

198100952
Numeric

Distinct count	6
Unique (%)	0.2%
Missing (%)	0.0%
Missing (n)	0
Infinite (%)	0.0%
Infinite (n)	0

Mean	0.0076336
Minimum	0
Maximum	7
Zeros (%)	99.7%

Toggle details

BBU-BoardType
Categorical

Distinct count	890
Unique (%)	29.5%
Missing (%)	0.0%
Missing (n)	0

1093-BBU	11
1118-SA	11
1141-SA	11
Other values (887)	2980

Toggle details

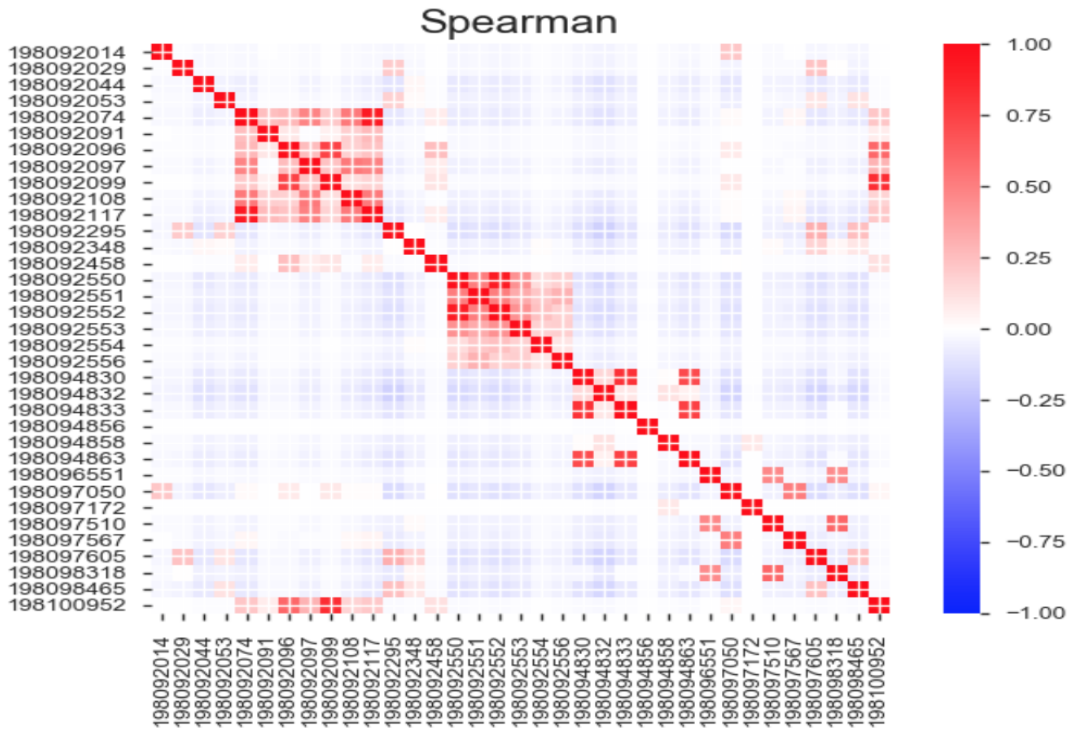
RaisedTimeAnioMes
Categorical

Distinct count	18
Unique (%)	0.6%
Missing (%)	0.0%
Missing (n)	0

Mar-2019	723
Jul-2019	341
Sep-2019	333
Other values (15)	1616

Toggle details

Correlations



Modelo Predicción Averías

SUMADOCASI

Modelo de prediccion de averias

Master en Business Intelligence y Big Data (2018/2019)

In [26]:
```
# Librerias de manipulacion de datos
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Librerias de modelados
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.neighbors import KNeighborsClassifier

# Librerias de analisis
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
```

CARGA DE DATOS

In [27]:
```
# Carga de informacion
df_train = pd.read_csv("WINDTELECOM_Averia_Dataset_Clustering.csv")
df_process = pd.read_csv("WINDTELECOM_Averia_Dataset_Clustering_September.csv")
```

In [28]:
```
# Revision de dataset
df_train.head()
```

Out[28]:

	BBU-BoardType	BoardType	RaisedTimeAnioMes	198092014	198092029	198092044	198092053	198092074	198092091	198092096	...	198097
0	1088-SA	SA	Oct-19	0	0	0	0	0	0	0	...	
1	1139-SA	SA	Oct-19	0	0	0	0	0	0	0	...	
2	1066-SA	SA	Oct-19	0	0	0	0	0	0	0	...	
3	1012-CCC	CCC	Oct-19	0	0	0	0	0	0	0	...	
4	1012-R8978 S2600L	R8978 S2600L	Oct-19	0	0	0	3	0	0	0	...	

5 rows × 40 columns

In [29]:
```
df_train.tail()
```

Out[29]:

	BBU-BoardType	BoardType	RaisedTimeAnioMes	198092014	198092029	198092044	198092053	198092074	198092091	198092096	...	198
3008	1352-BBU	BBU	Jul-18	0	1	0	0	0	0	0	...	
3009	1038-CCC	CCC	May-18	0	0	0	0	0	0	0	...	
3010	1226-CCC	CCC	May-18	0	0	0	0	0	0	0	...	
3011	1318-BBU	BBU	May-18	0	1	0	0	0	0	0	...	
3012	1207-CCC	CCC	Jan-18	0	0	0	0	0	0	0	...	

5 rows × 40 columns

ANALISIS DE DATOS

In [30]:
```
# verificar cantidad de datos que hay en los dataset
print(df_train.shape)
```

```
(3013, 40)
```

In [31]:
```
# verificar el tipo de datos conenida en ambos dataset
print(df_train.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3013 entries, 0 to 3012
Data columns (total 40 columns):
BBU-BoardType         3013 non-null object
BoardType             3013 non-null object
RaisedTimeAnioMes     3013 non-null object
198092014             3013 non-null int64
198092029             3013 non-null int64
198092044             3013 non-null int64
198092053             3013 non-null int64
198092074             3013 non-null int64
198092091             3013 non-null int64
198092096             3013 non-null int64
```

In [12]:
```
# estadistica de los dataset
print(df_train.describe())
```

```
          198092014     198092029     198092044     198092053     198092074  \
count  3013.000000   3013.000000   3013.000000   3013.000000   3013.000000
mean      0.012944      0.025888      0.282443      0.744109      0.306671
std       0.150800      0.248489      2.936504     10.164316      2.496654
min       0.000000      0.000000      0.000000      0.000000      0.000000
25%       0.000000      0.000000      0.000000      0.000000      0.000000
50%       0.000000      0.000000      0.000000      0.000000      0.000000
75%       0.000000      0.000000      0.000000      0.000000      0.000000
max       3.000000      4.000000     98.000000    385.000000     66.000000

          198092091     198092096     198092097     198092099     198092108  ...  \
count  3013.000000   3013.000000   3013.000000   3013.000000   3013.000000  ...
mean      0.059409      0.027215      0.353136      0.015931      0.128112  ...
std       1.048043      0.492227      3.090719      0.352959      0.986050  ...
min       0.000000      0.000000      0.000000      0.000000      0.000000  ...
25%       0.000000      0.000000      0.000000      0.000000      0.000000  ...
50%       0.000000      0.000000      0.000000      0.000000      0.000000  ...
75%       0.000000      0.000000      0.000000      0.000000      0.000000  ...
max      30.000000     21.000000     51.000000     16.000000     17.000000  ...

          198096551     198097050     198097172     198097510     198097567  \
count  3013.000000   3013.000000   3013.000000   3013.000000   3013.000000
mean      0.018918      0.491869      0.000332      0.077000      0.094590
std       0.172784      3.449168      0.018218      0.495711      0.506938
min       0.000000      0.000000      0.000000      0.000000      0.000000
25%       0.000000      0.000000      0.000000      0.000000      0.000000
50%       0.000000      0.000000      0.000000      0.000000      0.000000
```

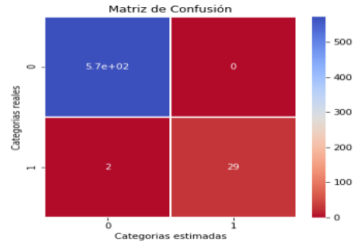
MODELOS Y MATRIZ DE CONFUSION

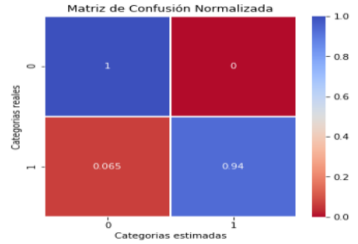
In [13]:
```
# metodos generales

def mostrar_matriz_confusion(y_test, pred_y):
    conf_matrix = confusion_matrix(y_test, pred_y)

    fig = plt.figure(figsize=(12,5))
    ax = fig.add_subplot(1,2,1)

    sns.heatmap(conf_matrix,cmap='coolwarm_r',linewidths=0.5,annot=True,ax=ax)
    plt.title('Matriz de Confusión')
    plt.ylabel('Categorias reales')
    plt.xlabel('Categorias estimadas')
    plt.show()

    conf_matrix_norm = confusion_matrix(y_test, pred_y)
    conf_matrix_norm = conf_matrix_norm.astype('float') / conf_matrix_norm.sum(axis=1)[:, np.newaxis]

    fig = plt.figure(figsize=(12,5))
    ax = fig.add_subplot(1,2,2)

    sns.heatmap(conf_matrix_norm,cmap='coolwarm_r',linewidths=0.5,annot=True,ax=ax)
    plt.title('Matriz de Confusión Normalizada')
    plt.ylabel('Categorias reales')
    plt.xlabel('Categorias estimadas')
    plt.show()

    print (classification_report(y_test, pred_y))
```

In [14]:
```
# Formateando las variables para los modelos

X = np.array(df_train.drop(['Averias','RaisedTimeAnioMes','BBU-BoardType','BoardType','Unificador'], 1))
y = np.array(df_train['Averias'])

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
```

In [15]:
```
# Implementando modelo de de clasificacion LogisticRegression
logreg = LogisticRegression()
logreg.fit(X_train, y_train)
Y_pred = logreg.predict(X_test)
print('Precisión Regresión Logistica:')
print(logreg.score(X_train, y_train))

mostrar_matriz_confusion(y_test, Y_pred)
```

```
//anaconda3/lib/python3.7/site-packages/sklearn/linear_model/logistic.py:432: FutureWarning: Default solver will be c
hanged to 'lbfgs' in 0.22. Specify a solver to silence this warning.
  FutureWarning)

Precisión Regresión Logística:
0.9962655601659751
```



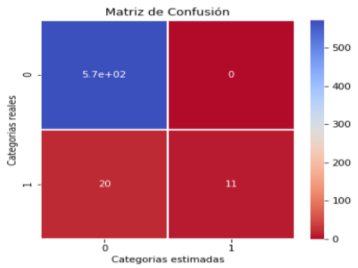


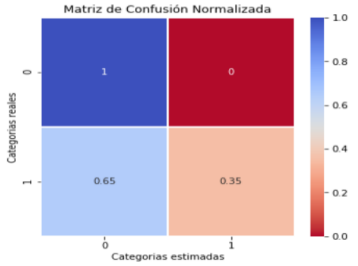

```
              precision    recall  f1-score   support

           0       1.00      1.00      1.00       572
           1       1.00      0.94      0.97        31

    accuracy                           1.00       603
   macro avg       1.00      0.97      0.98       603
weighted avg       1.00      1.00      1.00       603
```

In [16]:
```
# Implementando modelo de de clasificacion Support Vector Machines
svc = SVC()
svc.fit(X_train, y_train)
Y_pred = svc.predict(X_test)
print('Precisión Soporte de Vectores:')
print(svc.score(X_train, y_train))

mostrar_matriz_confusion(y_test, Y_pred)
```

```
//anaconda3/lib/python3.7/site-packages/sklearn/svm/base.py:193: FutureWarning: The default value of gamma will chang
e from 'auto' to 'scale' in version 0.22 to account better for unscaled features. Set gamma explicitly to 'auto' or
'scale' to avoid this warning.
  "avoid this warning.", FutureWarning)
```

```
Precisión Soporte de Vectores:
0.9975103734439834
```

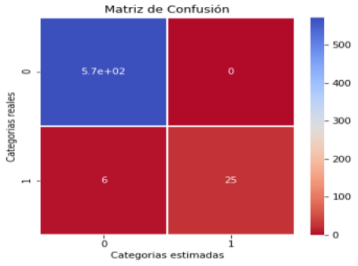


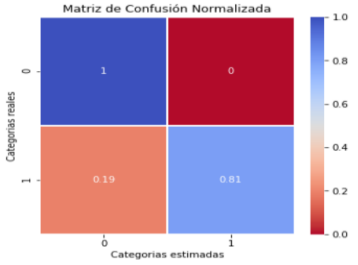


	precision	recall	f1-score	support
0	0.97	1.00	0.98	572
1	1.00	0.35	0.52	31
accuracy			0.97	603
macro avg	0.98	0.68	0.75	603
weighted avg	0.97	0.97	0.96	603

In [17]:
```
# Implementando modelo de de clasificacion KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors = 2)
knn.fit(X_train, y_train)
Y_pred = knn.predict(X_test)
print('Precisión Vecinos más Cercanos:')
print(knn.score(X_train, y_train))

mostrar_matriz_confusion(y_test, Y_pred)
```

```
Precisión Vecinos más Cercanos:
0.9929460580912863
```





	precision	recall	f1-score	support
0	0.99	1.00	0.99	572
1	1.00	0.81	0.89	31
accuracy			0.99	603
macro avg	0.99	0.90	0.94	603
weighted avg	0.99	0.99	0.99	603

OUTPUT

In [23]:
```
BBU = df_process['BBU-BoardType']
Type = df_process['BoardType']
df_predict = df_process.drop(['RaisedTimeAnioMes','BBU-BoardType','BoardType','Unificador'], 1)

prediccion_logreg = logreg.predict(df_predict)
out_logreg = pd.DataFrame({ 'BBU-BoardType' : BBU, 'BoardType' : Type,  'Averias': prediccion_logreg })
print('Predicción Regresión Logística:')
print(out_logreg.head())
```

```
Predicción Regresión Logística:
       BBU-BoardType     BoardType  Survived
0          1352-CCE1          CCE1         0
1          1316-CCE1          CCE1         0
2  1280-R8978 S2600L  R8978 S2600L         1
3           1280-CCC           CCC         1
4           1278-CCC           CCC         1
```

In [25]:
```
# Descarga de informacion predecida
export_csv = out_logreg.to_csv(r'C:\Users\admin\Desktop\SUMADOCASI_MODEL_CLUSTERING.csv',
                     index = None, header=True)
```

In [ ]:
```

```

Análisis de Supervivencia de Elementos BBU

SUMADOCASI

Modelo de supervivencia de averias

Master en Business Intelligence y Big Data (2018/2019)

In [48]:
```
%matplotlib inline

# Librerias de manipulacion de datos
import pandas as pd
import numpy as np

# Librerias de impresion
import matplotlib.pyplot as plt
import seaborn as sns

# Librerias de modelados
from lifelines import KaplanMeierFitter

# set some plotting aesthetics, similar to ggplot
# sns.set(palette = "colorblind", font_scale = 1.35, rc = {"figure.figsize": (12,9), "axe
s.facecolor": ".92"})
```

CARGA DE DATOS

In [49]:
```
# Carga de informacion
df_train = pd.read_csv("WINDTELECOM_Averia_Dataset_Survivel.csv")
df_process = pd.read_csv("WINDTELECOM_Averia_Dataset_Survivel_September.csv")
```

In [50]:
```
# Revision de dataset
df_train.head()
```

Out[50]:

	BBU-BoardType	BoardType	Duracion	Averias
0	1001-BBU	BBU	3	1
1	1001-BPL1	BPL1	3	0
2	1001-CCC	CCC	3	0
3	1001-FCE	FCE	3	0
4	1001-R8978 S2600L	R8978 S2600L	3	0

In [51]:
```
df_train.tail()
```

Out[51]:

	BBU-BoardType	BoardType	Duracion	Averias
885	3087-CCE1	CCE1	3	1
886	3087-R8964 S2600	Otros	3	1
887	3087-R8978 S2600L	R8978 S2600L	3	1
888	9005-B33TRA26A	Otros	3	1
889	9011-BBU	BBU	3	0

In [52]:
```
# verificar cantidad de datos que hay en los dataset
print(df_train.shape)
```

```
(890, 4)
```

In [53]:
```
# verificar el tipo de datos conenida en ambos dataset
print(df_train.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 890 entries, 0 to 889
Data columns (total 4 columns):
BBU-BoardType    890 non-null object
BoardType        890 non-null object
Duracion         890 non-null int64
Averias          890 non-null int64
dtypes: int64(2), object(2)
memory usage: 27.9+ KB
None
```

In [54]:
```
# estadistica de los dataset
print(df_train.describe())
```

```
         Duracion     Averias
count  890.000000  890.000000
mean     2.888764    0.244944
std      1.298952    0.430296
min      1.000000    0.000000
25%      2.000000    0.000000
50%      3.000000    0.000000
75%      4.000000    0.000000
max      6.000000    1.000000
```

MODELOS

In [74]:
```
BBUTypes = ["SA", "BBU", "BPL1",
            "CCC", "FCE", "R8978 S2600L",
            "BPN2", "CCE1"]

KmfBT  = {}

# Implementando modelo de supervivencia
for bt in BBUTypes:

    print(bt)
    df_selected = df_train[df_train.BoardType == bt]

    kmf = KaplanMeierFitter()
    kmf.fit(durations = df_selected.Duracion, event_observed = df_selected.Averias)

    predict = kmf.predict([2,3,5,6,10])

    KmfBT[bt] = kmf

    print(predict)
    print('')
```

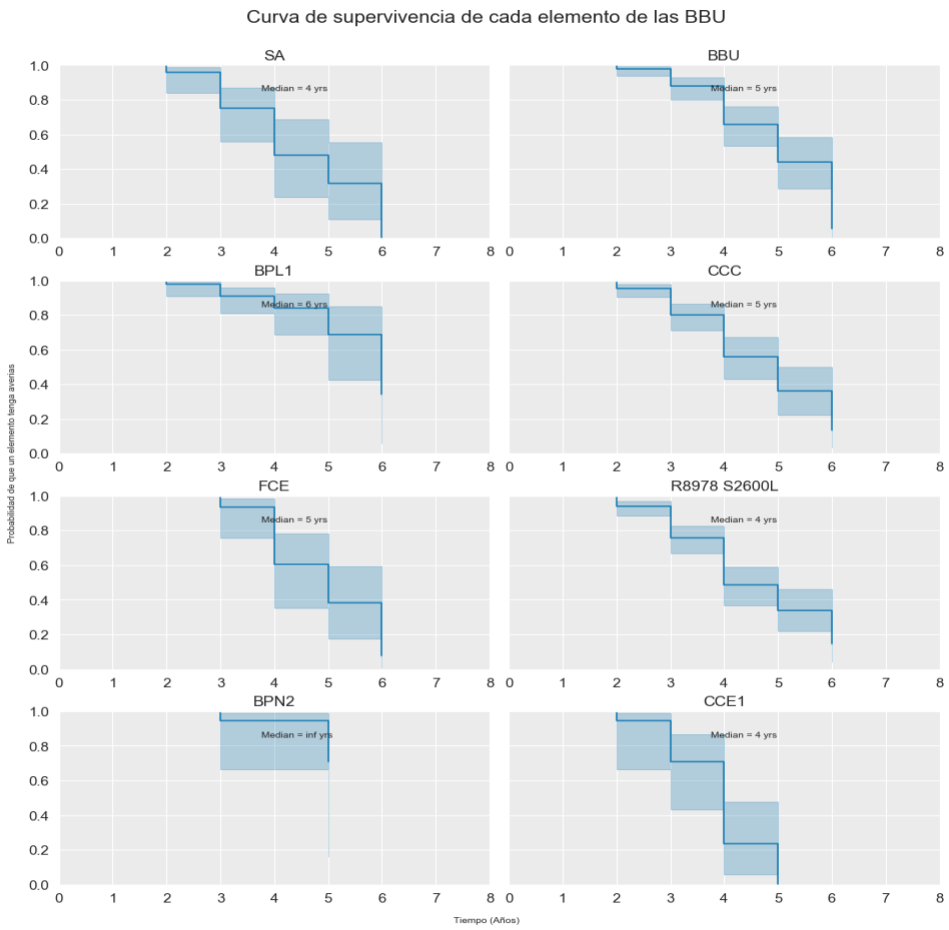
```
SA
2     0.956522
3     0.751553
5     0.318841
6     0.000000
10    0.000000
Name: KM_estimate, dtype: float64

BBU
2     0.979730
3     0.878060
5     0.439030
6     0.054879
10    0.054879
Name: KM_estimate, dtype: float64
```

```
R8978 S2600L
2     0.937500
3     0.755461
5     0.338447
6     0.145049
10    0.145049
Name: KM_estimate, dtype: float64

BPN2
2     1.000000
3     0.944444
5     0.708333
6     0.708333
10    0.708333
Name: KM_estimate, dtype: float64

CCE1
2     0.944444
3     0.708333
5     0.000000
6     0.000000
10    0.000000
Name: KM_estimate, dtype: float64
```

ANALISIS

In [75]:

```
# Impresion de analisis por
BBUTypes = ["SA", "BBU", "BPL1",
             "CCC", "FCE", "R8978 S2600L",
             "BPN2", "CCE1"]


fig, axes = plt.subplots(nrows = 4, ncols = 2, sharey = True,
                         figsize=(13,15))

for pos, ax in zip(BBUTypes, axes.flatten()):
    d = draft_df[draft_df.BoardType == pos]
    kmf_by_pos = KaplanMeierFitter()


    kmf_by_pos.fit(durations = d.Duracion, event_observed = d.Averias)
    kmf_by_pos.plot(ax=ax, legend=False)

    ax.annotate("Median = {:.0f} yrs".format(kmf_by_pos.median_), xy = (.47, .85),
                xycoords = "axes fraction")

    ax.set_xlabel("")

    ax.set_title(pos)

    ax.set_xlim(0,8)
    ax.set_ylim(0,1)


fig.tight_layout()

fig.text(0.5, -0.01, "Tiempo (Años)", ha="center")

fig.text(-0.01, 0.5, "Probabilidad de que un elemento tenga averias",
         va="center", rotation="vertical")

fig.suptitle("Curva de supervivencia de cada elemento de las BBU",
             fontsize=20)
fig.subplots_adjust(top=0.92)

plt.show()
```



OUTPUT

In [92]:
```
OutputData = []

for index, row in df_process.iterrows():

    try:
        BBU = row['BBU-BoardType']
        BType = row['BoardType']
        Averias = row['Averias']
        D = row['Duracion']
        Pred = KmfBT[BType].predict(D) * 100

        OutputData.append((BBU,BType,Averias,D,Pred))
    except:
        continue


dfObj = pd.DataFrame(OutputData, columns = ['BBU-BoardType' , 'BoardType', 'Averias','Dur
acion' , 'Supervivencia'])

dfObj.head()
```

Out[92]:

	BBU-BoardType	BoardType	Averias	Duracion	Supervivencia
0	1280-R8978 S2600L	R8978 S2600L	1	4	48.349515
1	1280-CCC	CCC	1	4	55.933101
2	1278-CCC	CCC	1	2	95.588235
3	1276-CCC	CCC	1	2	95.588235
4	1274-CCC	CCC	1	6	13.633693

In [93]:
```
# Descarga de informacion predecida
export_csv = dfObj.to_csv(r'C:\Users\admin\Desktop\SUMADOCASI_MODEL_SURVIVEL.csv',
                          index = None, header=True)
```

In [ ]: