ANÁLISIS Y MODELAMIENTO DEL SISTEMA DE LÍNEAS DE ESPERA DEL CENTRO DE SERVICIOS EN LA COMPAÑÍA IBM PARA LA REDUCCIÓN DEL ÍNDICE DE ABANDONO DE LLAMADAS

FABIAN LEONARDO DUEÑAS SUÁREZ

FACULTAD DE INGENIERÍA ELECTRÓNICA
UNIVERSIDAD SANTO TOMÁS
BOGOTÁ
2017

ANÁLISIS Y MODELAMIENTO DEL SISTEMA DE LÍNEAS DE ESPERA DEL CENTRO DE SERVICIOS EN LA COMPAÑÍA IBM PARA LA REDUCCIÓN DEL ÍNDICE DE ABANDONO DE LLAMADAS

FABIÁN LEONARDO DUEÑAS SUÁREZ

Monografía de pasantía empresarial presentada como requisito para optar al título de INGENIERO ELECTRÓNICO

Tutor de grado: EYBERTH ROLANDO ROJAS MARTÍNEZ
Magíster en Ingeniería Electrónica

FACULTAD DE INGENIERÍA ELECTRÓNICA
UNIVERSIDAD SANTO TOMÁS
BOGOTÁ
2017

CONTENIDO

Pág.
1. INTRODUCCIÓN 12

2. NOMBRE DEL PROYECTO....13

3. PROBLEMA 13

4. ANTECEDENTES 14
4.1. GENERAL MOTOR CORPORATION (GM)....14
4.2. KEYCORP 15
4.3. HEWLETT-PACKARD (HP)....16

5. JUSTIFICACIÓN 17

6. OBJETIVOS 18
6.1. OBJETIVO GENERAL....18
6.2. OBJETIVOS ESPECÍFICOS....18

7. FACTIBILIDAD 19

8. MARCO TEÓRICO....20
1.1 TEORÍA DE COLAS 20
1.2 RED NEURONAL 20
1.3 MATRIZ DE CONFUSIÓN....21

9. DISEÑO METODOLÓGICO....24

10. PLANEACIÓN Y DESARROLLO DEL PROYECTO 25

11. IDENTIFICACIÓN DE LAS NECESIDADES QUE DEBE SUPLIR LA MESA DE AYUDA 28

12. IDENTIFICACIÓN Y RECOLECCION DE INFORMACIÓN DE LOS FACTORES QUE EVALUAN EL DESEMPEÑO DE LA MESA DE AYUDA 31

12.1. INFRAESTRUCTURA PARA LA ATENCIÓN Y MONITORIZACIÓN DE LLAMADAS. 31
12.2. INFRAESTRUCUTRA TECNOLÓGICA SOFTWARE PARA LA CONSOLIDACIÓN Y CONTROL DE TICKETS. 32

13. TRATAMIENTO Y ANÁLISIS DE INFORMACIÓN 34
13.1. DISTRIBUIDOR AUTOMÁTICO DE LLAMADAS Y HERRAMIENTA DE GESTIÓN DE TICKETS 34
13.2. MODELAMIENTO DE LA MESA DE AYUDA Y SIMULACIÓN EN EL SOFTWARE ARENA 43

14. PRESENTACIÓN DEL PLAN DE MEJORA 46
14.1. SOLUCIÓN DE PROBLEMAS DE ENCOLAMIENTO MEDIANTE LA TEORÍA DE COLAS 46
14.2. DISEÑO E IMPLEMENTACIÓN DEL SOFTWARE DE DIAGNÓSTICO BASADO EN EL ENTRENAMIENTO DE REDES NEURONALES 51
14.3. DISEÑO E IMPLEMENTACIÓN DE CHATBOT PARA LA REDUCCIÓN DEL ENCOLAMAMIENTO EN LA MESA DE AYUDA. 56
14.4. ANÁLISIS DE LOS RESULTADOS FINALES MEDIANTE LA TEORIA DE COLAS Y SIMULACIÓN EN EL SOFTWARE ARENA 60

15. CONCLUSIONES. 65

16. BILIOGRAFÍA 66

ÍNDICE DE FIGURAS

Figura 1. Los 7 pasos para la mejora de los procesos.....26

Figura 2 Representación de la herramienta de gestión del ACD de AVAYA y el teléfono 1400 Series Digital Deskphones usado en la mesa de ayuda de IBM. 32

Figura 3. Diagrama de barras representando llamadas ofrecidas y contestadas vs días de la semana.....36

Figura 4. Diagrama de barras representando la tasa de disponibilidad con target señalado del 99.6%. 37

Figura 5. Diagrama de barras representando llamadas ofrecidas y contestadas en menos de 15 segundos vs días de la semana. 38

Figura 6. Diagrama de barras representando la tasa de velocidad de respuesta diaria con target señalado del 85%. 38

Figura 7. Diagrama de barras representando llamadas ofrecidas y no contestadas con tiempo de espera superior a 15 segundos vs días de la semana. 39

Figura 8. Diagrama de barras representando la tasa de abandono diaria con el target señalado del 5%. 39

Figura 9. Relación entre el tipo de servicio y tipo de contacto en la mesa de ayuda. 41

Figura 10. Representación en burbuja relacionando el TOP 15 por cantidad de tickets de tipos de plataforma intervenidos por la mesa de ayuda. 42

Figura 11. Representación de la cantidad de escalamientos vs prioridad según el tipo de contacto con la mesa de ayuda. 42

Figura 12. Simulación de la mesa de ayuda de IBM en software de simulación Arena. 44
Figura 13. Resultados del modelamiento de la mesa de ayuda de IBM con 12 agentes. 45
Figura 14. Comportamiento del costo total según la variación del número de servidores. 49
Figura 15. Variación Costo total vs cantidad de servidores, en la situación actual de la mesa de ayuda, basada en el costo mínimo de trabajo de los usuarios. 50
Figura 16.Variación Costo total vs cantidad de servidores, en la situación actual de la mesa de ayuda, basada en el costo máximo de trabajo de los usuarios. 51
Figura 17. Interfaz gráfica de software de diagnóstico de problemas basado en el entrenamiento de redes neuronales. 54
Figura 18. Diagnóstico final de software basado en entrenamiento de redes neuronales. 54
Figura 19. Mensaje de validación del software de diagnóstico de problemas para el reentrenamiento de la red neuronal 54
Figura 20. Ventana emergente de notificación de software indicando que una nueva red neuronal se ha entrenado 55
Figura 21. Diagrama de flujo de chatbot implementando Watson Conversation. ..59
Figura 22. Conversaciones en Chatbot de Watson Analytics mencionando problemas de VPN y solicitud de instalación de aplicativo 60

Figura 23. Variación Costo total vs cantidad de servidores, en la situación hipotética donde se implementase la solución, basada en el costo mínimo de trabajo de los usuarios. 61
Figura 24. Variación Costo total vs cantidad de servidores, en la situación hipotética donde se implementase la solución, basada en el costo máximo de trabajo de los usuarios. 62
Figura 25. Cantidad de llamadas atendidas y abandonas con 16 agentes, con software de simulación Arena. 62
Figura 26. Configuración de la tasa de entrada de llamadas en simulación en Arena, con los valores actuales de la mesa de ayuda 69
Figura 27. Configuración del sistema de distribución de usuario en software de simulación Arena. 69
Figura 28. Configuración de la tasa de atención de cada uno de los usuarios de la mesa de ayuda. 69
Figura 29. Configuración de la probabilidad de que una llamada sea solucionada por los diferentes grupos de trabajo en software Arena. 69
Figura 30. Reporte de resultados de la simulación de la mesa de ayuda. 70
Figura 31. Reporte de resultados de cada uno de los agentes de la mesa de ayuda en la simulación del software Arena. 70
Figura 32. Asignación de vectores de entrada y salida para el entrenamiento de la red neuornal 72
Figura 33. Configuración de la red neuronal en Matlab: número de neuronas ocultas, cantidad de iteraciones y tiempo de entrenamiento. 73

Figura 34. Cantidad de iteraciones vs media cuadrada del error; basada en el entrenamiento de la red neuronal. 73
Figura 35. Configuración de intentos en el flujo de conversación. 76
Figura 36. Configuración de entidades en el flujo de la conversación. 76
Figura 37. Resultados del modelamiento de la mesa de ayuda de IBM si se tuviesen 17 agentes. 77
Figura 38. Reporte de resultados de los tiempos de espera con la ayuda de 17 agentes en la mesa de ayuda. 77

ÍNDICE DE TABLAS

Tabla 1. Base de datos extraída de la central de distribución de llamadas AVAYA. 35

Tabla 2. Indicadores diarios de calidad de la mesa de ayuda. 36

Tabla 3. Promedio y varianza de llamadas entrantes a la mesa de ayuda según métricas mensuales. 37

Tabla 4. Cantidad de tickets ingresados a la mesa de ayuda según el tipo de ingreso. 40

Tabla 5. Cantidad de tickets según el tipo de servicio atendido telefónicamente por los agentes de le mesa de ayuda. 40

Tabla 6. Cantidad de tickets gestionados por grupo resolutor. 41

Tabla 7. Cantidad de tickets atendidos por los agentes de la mesa de ayuda en el mes de Enero.....43

Tabla 8. Cantidad de llamadas entrantes a la mesa de ayuda, tasa media diaria y tasa media por hora 43

Tabla 9. Cantidad de tickets por día y media de tickets por hora.....44

Tabla 10. Comparación de los indicadores de calidad de la mesa de ayuda, basado en la simulación Arena. 45

Tabla 11. Cantidad de tickets resueltos por la mesa de ayuda y el soporte en sitio según el tipo sistema operativo.....52

Tabla 12. Cantidad de tickets según el método de solución relacionado a problemas del sistema operativo. 53

Tabla 13.Cantidad de tickets según el Top de plataformas intervenidas por la mesa de ayuda y sitio. 57
Tabla 14. Cantidad de tickets según el método de solución empleado en la plataforma Internet Explorer.....58
Tabla 15. Cantidad de tickets según el método de solución empleado en la plataforma Outlook.....58
Tabla 16. Cantidad de tickets según el método de solución empleado en la plataforma Lync. 58
Tabla 17. Cantidad de tickets según el tipo de plataforma donde se optimizó la solución del chatbot y el software de diagnóstico de problemas 61
Tabla 18. Comparación de la totalidad de llamadas entrantes, media diaria y media por hora de la situación actual de la mesa de ayuda y su excepción según plan de mejora. 61
Tabla 19. Comparación de los indicadores de calidad de la mesa de ayuda con 16 agentes e implementando el chatbox y software de diagnóstico. 63
Tabla 20. Comparación de los indicadores de calidad de la mesa de ayuda en caso de implementar el chatbot de Watson y el software de diagnóstico, basado en la simulación Arena.....77

ÍNDICE DE ANEXOS

ANEXO A. Configuración y emisión de reportes en la simulación inicial con el software Arena. 69

ANEXO B. Código de programación en Matlab para calcular la optimización de costos en la mesa de ayuda mediante la teoría de colas 71

ANEXO C. Configuración de vectores de entrada y salida para el entrenamiento de la red neuronal. 72

ANEXO D. Código de programación en Matlab para entrenar la red neuronal. 72

ANEXO E. Resultados del entrenamiento de la red neuronal en Matlab. 73

ANEXO F. Código de programación para la interfaz GUI en Matlab 74

ANEXO G. Diseño, implementación y resultados del chatbot de Watson Conversation en la mesa de ayuda. 76

ANEXO H. Resultados finales implementando el Chatbox de Watson y el software de diagnóstico de problemas 77

1. INTRODUCCIÓN

Las mesas de ayuda son un punto de gestión central para el diagnóstico y solución de problemas, el servicio que ofrecen puede llegar a ser muy cuestionado ya que su diseño está enfocado en atender gran cantidad de usuarios y lograr su satisfacción, que puede no llegar a cumplirse debido a los largos de tiempos de espera o soluciones ineficientes.

Este estudio se encargará modelar un centro de servicios de IBM como una línea de espera con el objetivo de identificar los indicadores claves de desempeño para definir en primera instancia si la cantidad de personal dedicado es la más adecuada y lograr un balance entre los costos de la operación y la calidad del servicio.

Mediante la recolección de información dentro de los sistemas de medición de IBM se podrán extraer diferentes factores de rendimiento como el tiempo de espera del usuario, los tiempos de atención de cada uno de los agentes o identificar qué problemas pueden presentarse con mayor recurrencia y determinar diferentes planes de acción.

Con el objetivo de mitigar los encolamientos presentados en el sistema, se diseñará un software que evitará que los usuarios entren a la línea de espera, donde se les prestará un servicio de diagnóstico y solución capaz de suplir sus necesidades. Dado que en el ámbito de las “Tecnologías de la Información” la resolución de problemas varía constantemente será objeto de estudio el entrenamiento de una red neuronal que se adapte a la retroalimentación constante de los usuarios.

Cotidianamente los usuarios no solamente se contactan al centro de servicios para que brinden un diagnóstico técnico sino también a pedir información o indicaciones de un determinado procedimiento, situación puede ser solucionada mediante la implementación de un chat automatizado que igualmente sea programado para suplir las necesidades del negocio y así mismo disminuir la cantidad de llamadas al sistema.

Mediante la implementación de un sistema de gestión logístico y financiero más eficiente y el desarrollo del software basado en técnicas de inteligencia artificial, se busca finalmente el bienestar del trabajador, la satisfacción del usuario final y el valor agregado al negocio.

2. NOMBRE DEL PROYECTO

Análisis y modelamiento del sistema de líneas de espera del centro de servicios en la compañía IBM para la reducción del índice de abandono de llamadas.

3. PROBLEMA

El centro de servicios de IBM se encarga de darle soporte técnico a la infraestructura IT (Tecnologías de la información) de las empresas que requieran adquirir el servicio. Para poder suplir las necesidades del cliente, cuenta con una mesa de ayuda, la cual diariamente recibe en promedio 300 llamadas, pero cerca del 85% es recibida, las restantes son llamadas abandonadas por el usuario, debido a tiempos de espera prolongados ocasionados por encolamientos en el sistema.

El aumento de la cantidad de llamadas puede ser ocasionado por una falla masiva que presenten los usuarios al ingresar a un aplicativo en línea (alojado en la internet, intranet o extranet) debido a problemas en los servidores, dichas incidencias deben ser solucionadas por equipos de trabajo o proveedores especializados.

Existe otro tipo de falla que afecta directamente la estación de trabajo del usuario (donde puede haber un computador, una impresora, entre otros) y que podría ser solucionada por una persona que posea un dominio básico de informática, pero dada la situación actual, gran cantidad de usuarios se comunican con la mesa de ayuda, para que un agente resolutor de problemas solucione su inconveniente, el tiempo de atención que tarda puede ser demasiado largo, genere la desatención de múltiples usuarios y así mismo encolamientos.

Para mitigar el problema será necesario analizar si la cantidad de agentes en la mesa de ayuda es la adecuada y si los incidentes presentados en la estación del trabajo pueden ser atendidos por un software de diagnóstico.

4. ANTECEDENTES

La teoría de colas inició a principios del siglo XIX en el campo de la ingeniería telefónica, cuyo precursor fue A.K. Erlang, un empleado de la Danosh Telephone Company en Copenhague).[1] Hasta el día de hoy se han presentado múltiples proyectos basadas en dicho planteamiento para la optimización de costos y tiempos de operación dentro de la industria. Algunos proyectos se presentan a continuación.

4.1. GENERAL MOTOR CORPORATION (GM)

Por muchas décadas, la compañía gozó de tener la posición de líder mundial en la industria automotriz antes de que Toyota ocupara dicho lugar. GM tiene operaciones de manufactura en 32 países, emplea a más de 300,000 personas a nivel mundial y genera ganancias anuales cercanas a los 200,000 millones de dólares. Sin embargo, desde finales de los ochenta, cuando la productividad de las plantas de GM ocupaba un lugar muy por debajo de la industria, la posición en el mercado de la compañía se ha deteriorado debido a la creciente competencia extranjera.

Para contrarrestar los efectos de dicha competencia, la alta administración de GM inicio un proyecto de investigación de operaciones a largo plazo con el fin de predecir y mejorar la producción en las cientos de líneas de montaje de la compañía en todo el mundo. El objetivo fue incrementar en gran medida la productividad de sus operaciones de manufactura y, por ende, lograr una ventaja estratégica competitiva.

La herramienta analítica más importante que se utilizó en este proyecto ha sido un modelo de colas complejo que utiliza un simple modelo de un solo servidor. El modelo consta de una línea de producción de dos estaciones, cada una de las cuales se modela como un sistema de colas de un solo servidor con tiempos constantes entre llegadas y tiempos constantes de servicio.

La aplicación de la teoría de colas, junto con los sistemas de recolección de datos que la soportan, ha dado enormes beneficios a GM. De acuerdo con fuentes imparciales de la industria, sus plantas, que una vez fueron de las menos productivas en la industria, ahora se encuentran entre las más altas. Las mejoras resultantes en cuanto a producción en más de 30 plantas de automóviles y 10 países han representado una ganancia de 2,100 millones de dólares en ahorros y ganancias comprobados. Estos dramáticos resultados llevaron a General Motors a ganar en 2005 el primer lugar en la competencia internacional por el Premio Franz

[1] HILLIER, Frederick S; LIEBERMAN, Gerald J. Introducción a la investigación de operaciones. 9 ed. Mexico DF: McGraw-Hill/Interamericana editores, 2010. p. 750.

Edelman al desempeño en investigación de operaciones y ciencias de la administración.[2]

4.2. KEYCORP

Es una compañía cuya casa matriz está en Cleveland, Ohio, se encuentra dentro de las 500 compañías de Fortune. Es la decimotercera compañía bancaria más grande de Estados Unidos. Emplea a 19,000 personas, tiene activos por 93,000 millones de dólares y ganancias anuales por 6,700 millones de dólares. La compañía se enfoca en prestar servicios bancarios a los consumidores y cuenta con 2.4 millones de clientes repartidos en 1,300 sucursales y oficinas afiliadas.

Para impulsar el crecimiento de sus negocios, la gerencia de KeyCorp puso en marcha un amplio estudio de investigaciones en las operaciones para determinar la forma de mejorar el servicio a los clientes (el cual se define como la reducción del tiempo de espera para recibir el servicio), y a la vez conservar los costos de personal en un rango moderado. Se fijó un objetivo en cuanto a la calidad del servicio que consistía en que el tiempo de espera de los clientes debería ser menor a 5 minutos.

La herramienta clave para analizar este problema fue el modelo de colas M/M/s (tasa de entrada y de atención con distribución de probabilidad exponencial y cantidad de servidores *s*), el cual demostró ser el más conveniente para esta aplicación. Para aplicarlo, se solicitaron datos que revelaron que el tiempo de servicio promedio que se requería para procesar un cliente era muy elevado, del orden de 246 segundos. Con este tiempo promedio de servicio y tasa de llegada en la media usualmente, el modelo indicó que era necesario incrementar el 30% el número de personas que brindan atención a los clientes con el fin de cumplir con el objetivo de calidad del servicio. Esta opción extremadamente costosa llevo a la alta gerencia a concluir que era necesario llevar a cabo una amplia campaña para reducir de manera drástica el tiempo de servicio promedio mediante la reingeniería de las sesiones con los clientes y del fomento de una mejor administración del personal. En un periodo de tres años, esta campaña dio como resultado una reducción del tiempo de servicio promedio hasta llegar a 115 segundos. La aplicación frecuente del modelo M/M/s revelo la forma en que la meta de calidad del servicio puede sobrepasarse de manera significativa, a la vez que se pueden reducir los niveles de personal mediante una programación optimizada de los empleados en las diferentes sucursales del banco.

El resultado neto se ha reflejado en ahorros de casi 20 millones de dólares por año con un servicio significativamente mejorado que permite que 96% de los clientes tengan que esperar menos de 5 minutos. Esta mejora se extendió a toda la compañía ya que el porcentaje de sucursales que cumple el objetivo de calidad de

[2] HILLIER, Frederick S; LIEBERMAN, Gerald J. Op. cit., p. 727

servicio aumento de 42 a 94 por ciento. Diferentes estudios también confirman un incremento significativo en la satisfacción del cliente.[3]

4.3. HEWLETT-PACKARD (HP)

Es una manufacturera de equipos electrónicos multinacional líder en su rama. En 1993 la compañía instalo un sistema de línea de ensamble mecanizado para fabricar impresoras de inyección de tinta en su planta de Vancouver, Washington, para satisfacer la explosiva demanda de este tipo de impresoras. Pronto se hizo evidente que el sistema instalado no sería tan rápido ni tan confiable como para satisfacer las metas de producción de la compañía. En consecuencia, se formó un equipo conjunto de científicos de la administración de HP y del Massachusetts Institute of Technology (MIT) para estudiar como rediseñar el sistema para mejorar su desempeño. El equipo de HP/MIT se percató con rapidez de que el sistema de línea de ensamble se podría modelar como un tipo especial de sistema de colas donde los clientes (las impresoras que debían ensamblarse) pasarían a través de una serie de servidores (operaciones de ensamblaje) en una secuencia fija.

Un modelo de colas especial para este tipo de sistema proporcionó con rapidez los resultados analíticos que se necesitaban para determinar el modo en que el sistema debería rediseñarse para lograr la capacidad que se requería de la forma más económica. Los cambios incluyeron la incorporación de espacios de almacenamiento en puntos estratégicos para mantener de mejor manera el flujo de trabajo hacia las estaciones subsecuentes y para contrarrestar el efecto de la descompostura de las maquinas. El nuevo diseño incremento alrededor de 50% la productividad y produjo aumentos en las ganancias de aproximadamente 280 millones de dólares por ventas de impresoras, así como utilidades adicionales por accesorios como productos. Esta aplicación innovadora del modelo de colas especial también le dio a HP un nuevo método para crear diseños de sistemas rápidos y eficaces en otras áreas de la compañía.[4]

[3] HILLIER, Frederick S; LIEBERMAN, Gerald J. Op. cit., p. 727
[4] Ibid. p. 715

5. JUSTIFICACIÓN

La implementación del modelamiento matemático mediante teoría de colas puede brindar soluciones, tanto para IBM que es la prestadora del soporte como para la empresa que solicita el servicio, debido a que cada tarea que se está dejando de realizar puede impedir el continuo flujo de un proceso de vital importancia.

Para IBM puede resultar un problema, debido a la falta de aprovechamiento de sus recursos, que se puede contrarrestar disminuyendo el tiempo de atención y así generar una optimización en la operación y en los costos del servicio.

Implementando inteligencia computacional basada en redes neuronales se podría desarrollar una herramienta de consulta para los analistas y los usuarios. Este software podría entrenarse con la solución de los problemas que se vayan presentando cotidianamente y así construir una base de datos que mejore cada vez más su efectividad. Teniendo una interfaz gráfica y la programación de redes neuronales especializada en la solución de problemas, los agentes de la mesa de ayuda disminuirían los tiempos de atención consultado el aplicativo, adicionalmente los usuarios obtendrían un dictamen de la plataforma sin tener que entrar a la cola del sistema y de esta forma evitar encolamientos.

La implementación de la teoría de colas con la optimización de los costos de operación y la inteligencia computacional con el sistema de interacción hombre-maquina, son un valor agregado a la industria que podría ser vendido como un servicio adicional y así marcar un hito en la integración de soluciones basadas en la ingeniería computacional e industrial.

Desarrollar esta solución brinda la posibilidad de aprovechar el tiempo del personal técnico, ya que el objetivo final es lograr la distribución de una carga laboral equilibrada, que sería el resultado de reconocer las falencias que existen en la operación y la construcción de estrategias de trabajo en equipo o reclutamiento de nuevos aspirantes. Finalmente, la satisfacción del trabajador será de mayor grado, cumpliendo con el propósito final, que va en pro de mejorar notablemente la calidad de vida de los empleados, disminuyendo la cantidad de estrés laboral y promoviendo el relacionamiento con sus pares, estado que se verá reflejado directamente en los índices de evaluación de desempeño en la mesa de ayuda.

Una vez logrados los objetivos se evidenciará un mejor ambiente laboral y flujo de trabajo, también denominado sinergia, que ayudará al constante crecimiento de la unidad de negocios.

.

6. OBJETIVOS

6.1. OBJETIVO GENERAL

Modelar y validar el sistema de líneas de espera en la mesa de ayuda de la empresa IBM con el fin de disminuir el índice de abandono de llamadas y aumentar la tasa de disponibilidad del servicio.

6.2. OBJETIVOS ESPECÍFICOS

Diseñar e implementar un sistema de gestión que disminuya los costos del servicio y costos de espera, estableciendo una cantidad específica de servidores.

Implementar una aplicación que permita dar soluciones inmediatas a los usuarios sin ingresar a la cola del sistema.

Reducir los costos de la operación implementando un software de diagnóstico técnico que minimice los tiempos de solución.

Asignar una carga laboral equitativa, generando estrategias en la administración del servicio y herramientas de soporte que apoyen constantemente la operación, logrando un equilibrio entre la calidad del servicio, la capacidad laboral y la satisfacción del cliente.

7. FACTIBILIDAD

Para poder acceder a material académico relacionado al desarrollo del proyecto se puede acudir a recursos académicos los cuáles están disponibles en internet, en la biblioteca de la Universidad Santo Tomás, en bibliotecas públicas o en bibliotecas con la cuales se tenga convenio. Adicionalmente los temas tratados están incluidos en las materias del pensum académico y se podría acudir directamente a los docentes que posean conocimientos especializados en determinada área.

Las herramientas informáticas que se implementarán serán Matlab y Arena. Matlab será utilizado con fines académicos y avalado con las licencias que cuenta la universidad, una vez sea desarrollado el proyecto y presentado a IBM, se estudiará la viabilidad de la adquisición del producto. El segundo programa es una plataforma desarrollada para simular líneas de espera la cual será utilizada bajo la licencia de estudiante y la adquisición del producto igualmente será evaluada por IBM.

Para realizar el modelamiento de las líneas de espera es necesario acudir a las métricas mensuales que maneja IBM, las cuales miden el rendimiento y calidad de cada una de las áreas del Centro de Soporte Técnico. A la hora de adquirir tales recursos fue necesario argumentar los fines académicos para lo cual fueron requeridos, ya que toda la información debe ser confidencial según políticas internas de la compañía. Como respuesta a la solitud, IBM puso a disposición la base de datos donde se consolida la información de la operación diaria, allí se encuentran indicadores tales como la tasa de velocidad de respuesta, la tasa de abandono, la tasa de solución, la tasa de disponibilidad, entre otros indicadores, con los cuales se realizará el respectivo modelamiento y análisis matemático.

Con el ánimo de mejorar el servicio brindado al cliente, el coordinador de la mesa de ayuda está dispuesto a analizar la propuesta para que la operación fluya mejor y así mismo se evidencien las falencias más comunes. Adicionalmente se validará el punto de vista de los agentes resolutores para que en la implementación del plan de mejora se cuenta con su aval y así mismo genere su satisfacción al mejorar sus condiciones de trabajo. Como beneficiado final estaría el usuario ya que la solución busca reducir los tiempos de atención y velar por que su operatividad no se vea afectada.

8. MARCO TEÓRICO

1.1 TEORÍA DE COLAS

La teoría de colas es el estudio de la espera en las distintas modalidades. Utiliza los modelos de colas para representar los tipos de sistemas de líneas de espera que surgen en la práctica. Las fórmulas de cada modelo indican cual debe ser el desempeño del sistema correspondiente y señalan la cantidad promedio de espera que ocurrirá en diversas circunstancias.

Por lo tanto, estos modelos de líneas de espera son muy útiles para determinar cómo operar un sistema de colas de la manera más eficaz. Proporcionar demasiada capacidad de servicio para operar el sistema implica costos excesivos; pero si no se cuenta con suficiente capacidad de servicio surgen esperas excesivas con todas sus desafortunadas consecuencias. Los modelos permiten encontrar un balance adecuado entre el costo de servicio y la cantidad de espera.

Los clientes que requieren un servicio se generan en el tiempo en una fuente de entrada. Luego, entran al sistema y se unen a una cola. En determinado momento se selecciona un miembro de la cola para proporcionarle el servicio mediante alguna regla conocida como disciplina de la cola. Se lleva a cabo el servicio que el cliente requiere mediante un mecanismo de servicio, y después el cliente sale del sistema de colas.[5]

1.2 RED NEURONAL

Las redes neuronales artificiales se definen como sistemas de mapeos no lineales cuya estructura se basa en principios observados en los sistemas nerviosos de humanos y animales. Constan de un número grande de procesadores simples ligados por conexiones con pesos. Las unidades de procesamiento se denominan neuronas. Cada unidad recibe entradas de otros nodos y genera una salida simple escalar que depende de la información local disponible, guardada internamente o que llega a través de las conexiones con pesos. Pueden realizarse muchas funciones complejas dependiendo de las conexiones.[6]

Las neuronas artificiales simples fueron introducidas por McCulloch y Pitts en 1943. Una red neuronal se caracteriza por los siguientes elementos:

- Conjunto de unidades de procesamiento o neuronas.

- Un estado de activación para cada unidad, equivalente a la salida de la unidad.

[5] HILLIER, Frederick S; LIEBERMAN, Gerald J. Op. cit., p. 708.
[6] PONCE CRUZ, Pedro. Inteligencia Artificial con aplicaciones a la ingeniería. 1 ed. México D.F: Alfaomega Grupo Editor, 2010. p.203.

- Conexiones entre las unidades, generalmente definidas por un peso que determina el efecto de una señal de entrada en la unidad.

- Una regla de propagación, que determina la entrada efectiva de una unidad a partir de las entradas externas.

- Una función de activación que actualiza el nuevo nivel de activación basándose en la entrada efectiva y la activación anterior.

- Una entrada externa que corresponde a un término determinado como referencia (bias) para cada unidad.

- Un método para reunir la información, correspondiente a la regla del aprendizaje

- Un ambiente en el que el sistema va a operar, con señales de entrada e incluso señales de error.[7]

1.3 MATRIZ DE CONFUSIÓN

Es un diseño de tabla que permite evaluar el rendimiento de un algoritmo que se emplea en un aprendizaje supervisado. Cada columna de la matriz representa el número de predicciones de cada clase, mientras que cada fila representa cada clase real.

Para analizar una matriz de confusión de múltiples clases, se realizará un ejemplo. Asumiendo que se tiene un problema con 3 clases A, B y C se tiene la siguiente matriz de confusión.

Tabla 1. Matriz de confusión de tres clases.

		PREDECIDO		
		A	B	C
ACTUAL	A	TP_A	E_{AB}	E_{AC}
	B	E_{BA}	TP_B	E_{BC}
	C	E_{CA}	E_{CB}	TP_C

Fuente: Elaboración del autor

donde E_{XY} se lee: “Cantidad de intentos que se predijo Y siendo X el valor.
Se deben tener en cuenta los siguientes conceptos:

[7] PONCE CRUZ, Pedro. Inteligencia Artificial con aplicaciones a la ingeniería. Op. cit., p.203.

- FN, “False Negative”: El total de falsos negativos para una clase se obtiene sumando todos los valores correspondientes a la fila excluyendo el verdadero positivo (TP, “True Positive”).

- FP, “False Positive”: El total de falsos positivos para una clase se obtiene sumando los valores correspondientes a la columna excluyendo el verdadero positivo (TP, “True Positive”).

- TN, “True Negative”: El total de verdaderos negativos se obtiene sumando toda la matriz de confusión excepto la fila y la columna de dicha clase.

- TP, “True Positive”: Es la cantidad de aciertos que se tuvo en determinada clase.

Los indicadores de rendimiento de una matriz de confusión son los siguientes:

- Exactitud: Se halla calculando la suma de todas las clasificaciones que fueron correctas dividio por el número total de clasificaciones.

- Precisión: Es calculado como se indica a continuación:

$$Precisión = \frac{TP}{TP + FP}$$

Siendo así, se tiene el siguiente ejemplo:

$$Precisión\ de\ A = \frac{TP_A}{TP_A + E_{BA} + E_{CA}}$$

- Sensibilidad: Corresponde a los verdaderos positivos dentro de una clase.

$$Sensibilidad = \frac{TP}{TP + FN}$$

Siendo de esta forma, se tiene el siguiente ejemplo:

$$Sensibilidad\ de\ B = \frac{TP_B}{TP_B + E_{BA} + E_{BC}}$$

- Especificidad: Corresponden a los verdaderos negativos dentro de una determinada clase. Se calcula como se indica a continuación:

$$Especificidad = \frac{TN}{TN + FP}$$

y de esta forma, para la clase C se calcularía:

$$Especificidad\ para\ C = \frac{TN_C}{TN_C + E_{AC} + E_{BC}}$$ [8]

[8] OLD DOMINION UNIVERSITY. Evaluating a classification model – What does precision and recall tell me? [en linea], [revisado el 5 de Diciembre]. Disponible en internet: http://www.cs.odu.edu/~mukka/cs795sum10dm/Lecturenotes/Day4/recallprecision.pdf

9. DISEÑO METODOLÓGICO

Siendo IBM una compañía que garantiza la calidad en la prestación del servicio técnico, se cuenta con sistemas de medición de métricas los cuales son entregadas al director de cada mesa de ayuda con el fin de retroalimentar a los analistas y buscar mes a mes mejores soluciones para gestionar más eficazmente su operación. Finalmente, bajo algunas condiciones de privacidad y confidencialidad, IBM accedió a brindar las métricas, con el fin de incentivar la investigación en fines académicos. Una vez obtenida la información se procede a realizar el modelamiento de la línea de espera mediante la utilización de la teoría de colas para finalmente obtener un sistema que pueda ser simulado y optimizado mediante diferentes técnicas que serán detalladas en el desarrollo del proyecto.

Para construir una base de datos lo suficientemente completa para aplicar las diferentes técnicas de inteligencia computacional, se deben identificar los problemas que más llegan a la mesa de ayuda y para ello se acudirá a la base de datos de gestión de solicitudes donde se extraerán los problemas y soluciones más comunes.

Para lograr desarrollar el proyecto se subdividió en las fases que se muestran a continuación

Recolección de datos: Se solicitan métricas de calidad a la compañía y adicionalmente se realiza el análisis de la situación actual de la mesa de ayuda.

Análisis y tabulado de la información: Se identifican las principales causas-raíz del encolamiento mediante métodos gráficos y analíticos.

Modelamiento del sistema: Se aplica la teoría de colas que, mediante cálculos matemáticos, es capaz de modelar el comportamiento de la línea de espera.

Mejora del sistema: Se realiza el desarrollo de aplicaciones con la utilización de técnicas basadas en inteligencia computacional para implementar herramientas de consulta y solución de problemas, mitigando así los encolamientos que presenta la mesa de ayuda. En este punto se encuentra concentrado el mayor tiempo dedicado a la investigación, donde se consultarán todos los soportes académicos y bibliográficos, al igual que se acudirá a los docentes que cuentan con el conocimiento de los temas a trabajar en el proyecto.

Conclusiones: Discernir las soluciones más rentables y que mejoren la calidad del servicio.

10. PLANEACIÓN Y DESARROLLO DEL PROYECTO

El núcleo del servicio de IBM Colombia se basa en la venta de servidores a compañías de cualquier sector industrial, ya sean pequeñas empresas emergentes en el mercado, que quieran salvaguardar sus datos de manera segura o incluso empresas estatales que manejan información confidencial y requieren de un servicio especializado que pueda reaccionar ante cualquier eventualidad; tales como ataques informáticos, desastres naturales, accidentes sobre las máquinas o cualquier incidencia que pueda afectar el rendimiento de las máquinas. La compañía que tenga dentro su portafolio la prestación de servicios de soporte a infraestructura IT requerirá tener un equipo o unidad de trabajo dedicado a brindar el soporte post venta, para ello IBM Colombia cuenta con la unidad de negocios TSS (Technical Service Support) que cuenta también con un gran catálogo de servicios enfocado a solucionar incidencias que presenten los clientes con cualquier equipo. La unidad de negocios cuenta con dos grupos de trabajo que son MVS (Multi Vendor Services) y Logo, el primero se basa en el soporte de infraestructura IT, lo que quiere decir que puede brindar apoyo en la resolución de inconvenientes relacionado a cualquier dispositivo electrónico, como mouses, teclados, CPUs, pantallas, teléfonos, datafonos, cajeros, entre algunos equipos, cabe resaltar que puede ser de cualquier marca, no importa que IBM no haya sido la empresa manufacturera. El segundo equipo de trabajo llamado Logo, está enfocado a *Traditional Services,* categorización asignada por la empresa a la prestación del soporte a servidores que ha fabricado, para este tipo de productos la empresa tiene un amplio portafolio de servicios.

IBM es una multinacional que se rige bajo estándares internacional de calidad en la prestación del servicio tales como ISO20000[9], donde se encuentran manuales de buenas prácticas como lo es ITIL (Information Technology Infrastructure Library), que nació después de que la compañía “Central Computer and Telecommunications Agency” (CCTA) del gobierno británico consolidara todos sus procesos y fuese una referencia o modelo a seguir para todas las compañías que prestaran servicios de IT[10]. La unidad MVS sigue las buenas prácticas que se consolidan en los manuales de ITIL, en éste se consolidan 36 procesos y cuatro funciones que los divide en cinco etapas, las cuales son Estrategia, Diseño, Transición, Operación del Servicio y Mejora Continua del Servicio[11], cada una de éstas es un módulo completo de estudio y la compañía que lo requiera puede certificar al personal mediante un examen.

---

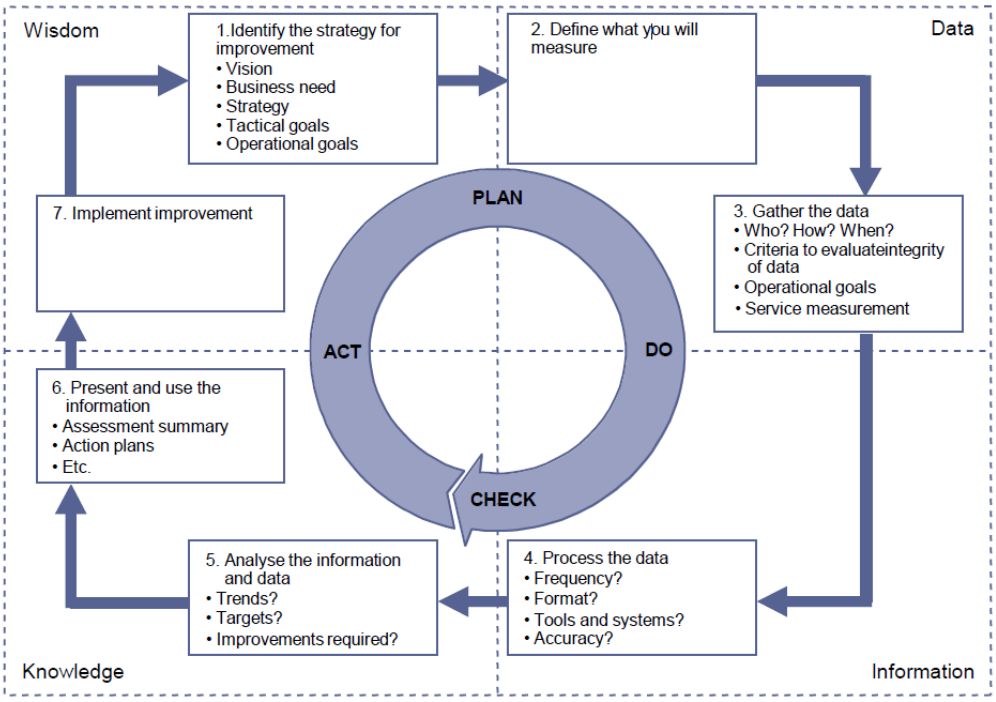
[9] IBM. IBM has obtained Corporate wide certifications for ISO 9001, ISO 14001, ISO 50001 and OHSAS 18001 [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www-935.ibm.com/services/us/en/it-services/iso-management-system-certifications.html.

[10] ITIL Central. In A Nutshell: A Short History of ITIL [en línea], [revisado de 2 Octubre de 2017]. Disponible en internet: http://itsm.fwtk.org/History.htm

[11] Haren, Van. Operación del Servicio, Basada en ITIL®V3. 1 ed. Van Haren Publishing, Zaltbommel, 2008. p.18.

La unidad de MVS de IBM Colombia en su operación diaria se rige bajo las buenas prácticas que se encuentra en el módulo de Operación del Servicio estipulado por ITIL, el cual cuenta con 5 procesos que son "*Gestión de Eventos, Gestión de Incidencias, Gestión de Requerimientos, Gestión de Problemas y Gestión de Accesos",* y cuatro funciones que son: "*Centro de servicios a usuarios (Help Desk, Services Desk), Gestión Técnica, Gestión de Operaciones I.T. y Gestión de aplicaciones*"[12]. El ciclo de vida de la prestación del servicio inicia en la mesa de ayuda (Help Desk) donde se realiza el primer contacto entre el proveedor de servicios y el usuario mediante una llamada telefónica, un chat o mediante una página web donde el catálogo de servicios es ofrecido al usuario. El objetivo de este trabajo es optimizar la prestación del servicio de la mesa de ayuda, basándose en el quinto módulo de buenas prácticas de ITIL llamado "Mejora Continua del servicio", que contiene siete pasos para poder lograrlo (el ciclo se visualiza en la Figura 1) y siendo enfocados a la problemática del caso serían los siguientes: Identificar las necesidades que debe suplir la mesa de ayuda, definir factores o métricas que incidan en el desempeño de la mesa de ayuda, recolectar la información necesaria, utilizar herramientas que ayuden a tratar la información, analizar el rendimiento, presentar planes de mejora e implementar la solución, cabe resaltar que el último paso no se vería reflejado en el trabajo y estaría sujeto a la decisión del gerente en la revisión del proyecto.

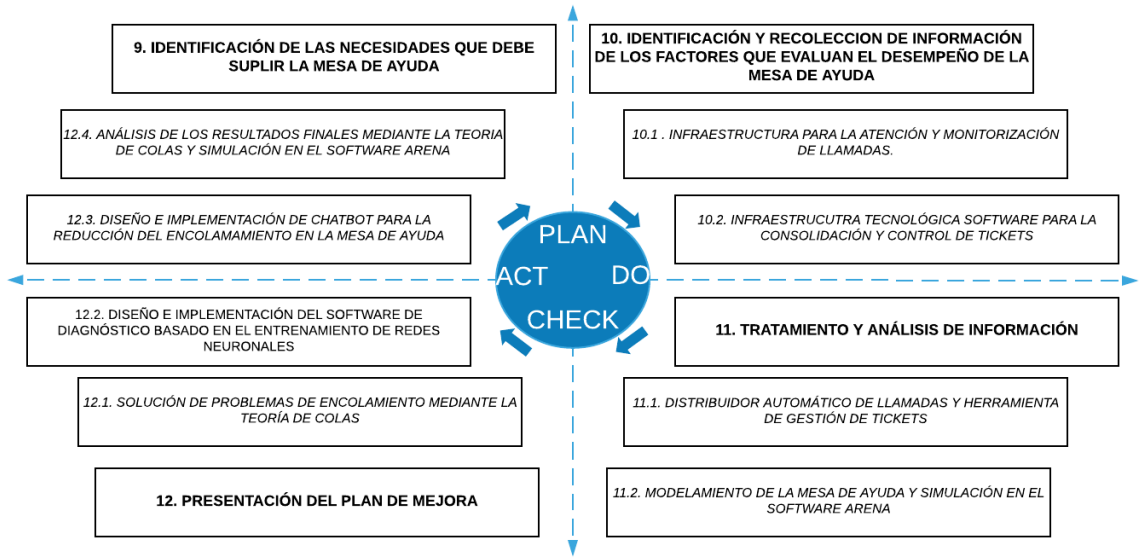
Figura 1. Los 7 pasos para la mejora de los procesos.



Fuente: IBM. Workshop - ITIL Foundations V3. Mexico DF: IBM Corporation, 2012.

[12] ITIL Central. In A Nutshell: The Five ITIL Volumes [en línea], [revisado de 2 Octubre de 2017]. Disponible en internet: http://itsm.fwtk.org/Contents.htm

Figura 2. Ciclo de mejora de Deming aplicado al desarrollo del proyecto



Fuente: Elaboración del autor.

En la Figura 2 se plantean las fases del desarrollo del proyecto basado en el ciclo de mejora de los procesos según las prácticas de ITIL. El primer paso será la identificación de las necesidades del cliente, posteriormente se recopilará la información que ayudará a evaluar la calidad del servicio, en esta fase se tratarán dos fuentes de información, la base de datos de la central de llamadas y la de la herramienta de creación y gestión de tickets.

Una vez obtenida la información, se realizará el tratamiento y análisis de datos, en este caso se evaluará cada métrica de desempeño para identificar las falencias más comunes en la mesa de ayuda y con ello sustentar los posibles planes de solución. Adicionalmente se modelará la mesa de ayuda y se usará un software de simulación, así se obtendrá un sistema que se comporte de forma semejante al sistema real, para modificar las diferentes variables y finalmente evaluar los resultados.

En el plan de mejora se implementarán dos tipos de solución, las que mitigan el encolamiento y las que optimizan los costos del servicio. Mediante la implementación del chatbot y software de diagnóstico mediante el entrenamiento de redes neuronales se planea reducir la cantidad de usuarios que ingresen al sistema y así evitar el desbordamiento de la línea de espera de la mesa de ayuda.

Para la optimización de los costos se utilizará la teoría de colas de Erlang, donde se hallará el modelo que permita disminuir los costos de servicio y de espera (costo de trabajo de los usuarios que estarían en cola) y finalmente hallar el capital de ahorro que se obtendría al implementar la solución.

11. IDENTIFICACIÓN DE LAS NECESIDADES QUE DEBE SUPLIR LA MESA DE AYUDA

La Mesa de Servicios se debe enfocar en buscar la concentración de las solicitudes de los usuarios finales en un punto único de contacto, facilitando una rápida y fluida comunicación, buscando reducir los tiempos de atención y solución de las incidencias o requerimientos reportados por los usuarios y brindando al cliente la posibilidad de interactuar con una única fuente de información a efectos del seguimiento de la prestación de los servicios de TI. Los servicios de IBM deberán estar basados en las mejores prácticas del mercado como son las descritas por ITIL (Infrastructure Technology Information Library).

Estas prácticas les permitirán a los usuarios recuperar de forma más rápida la disponibilidad de los servicios a través de la utilización de herramientas, procesos y control de eventos que habilitarán una continua y eficaz comunicación con el usuario, facilitándole el control sobre todos los servicios prestados. Los objetivos que se buscan con la implementación de una Mesa de Servicio con estas características son:

- Mejorar la percepción y satisfacción del servicio del área de tecnología.

- Dar al usuario final una mayor posibilidad de acceso al área de tecnología a través de un único punto de contacto.

- Contar con un modelo que permita dar respuesta oportuna a las solicitudes de servicio del usuario final.

Para que estos objetivos se cumplan es necesario que la Mesa de Servicio cuente con el conocimiento apropiado del negocio del cliente para canalizar en forma adecuada las necesidades de los usuarios finales, priorizando la atención según la urgencia e impacto del incidente o requerimiento generado.

La Mesa de servicio podría ser contactada por alguno de los siguientes medios:

- Acceso Telefónico: Esta línea de acceso es utilizada para aquellas solicitudes de servicio que requieren de atención inmediata o que para su ejecución no requieren autorización previa por parte de personal de TI del cliente.

- Correo Electrónico o Catálogo de Servicios de la Herramienta de Gestión: Se utiliza para aquellos requerimientos que por su naturaleza no necesiten de una solución o asesoría inmediata o que para su ejecución requieren de una aprobación previa por parte del área IT del cliente, por ejemplo: servicios tipo administración de usuarios, servicios que no han sido

contratados previamente o solicitudes de instalaciones, movimientos, adiciones, cambios y reclamos (IMAC).

- Chat**:** Esta línea de acceso será utilizada para el reporte de incidentes o solicitudes, generando un reporte en la herramienta de gestión para cada servicio atendido.

El usuario puede iniciar el contacto con la mesa de ayuda para que IBM realice gestión sobre cualquiera motivo referenciado a continuación:

- Incidentes: El soporte al usuario final del cliente está orientado a resolver las fallas que por diferentes causas los usuarios enfrenten en su sitio de trabajo. Comprende el soporte de primer nivel para el Software de usuario (sistema operativo, aplicaciones de oficina, aplicaciones específicas del cliente). Las fallas en el funcionamiento del Hardware reportado a la Mesa de Servicio son analizadas para tratar de proveer una solución de forma telefónica y si esto no es posible son enrutados al grupo de Soporte en Sitio correspondiente.

- Requerimientos: Son solicitudes de servicio relacionadas con el ambiente de microinformática y que no están originadas en fallas del ambiente IT de los usuarios finales. A modo de ejemplo: Un usuario solicita información sobre cómo ejecutar procedimientos asociados a las funciones del negocio que prestan las aplicaciones propietarias del cliente.

- Consultas: Son todas aquellas solicitudes de información que no están relacionadas con fallas o requerimientos sobre la infraestructura IT soportada; se refiere a los servicios que se prestan a través de la Mesa de Servicio. A este tipo de soporte igualmente se le creará un registro, se proveerá la información requerida cuando ésta se encuentre a disposición del personal del Mesa de Servicio o en su defecto se direccionará la consulta hacia el área que corresponda y se cerrará el registro de la actividad una vez atendida.

- Reclamos: Quejas de los usuarios sobre incidentes o requerimientos abiertos o cerrados.

- Llamadas escaladas a proveedores internos/externos del cliente: servicios sobre los cuales IBM no tiene responsabilidad.

Siendo identificadas las necesidades que debe suplir la mesa de ayuda se puede definir a continuación los roles y responsabilidades de cada de uno de los agentes resolutores:

- Abrir un ticket en la herramienta de Gestión “Máximo” (desarrollada por IBM), para cada contacto entre la Mesa de Ayuda y el usuario.

- Acceder a la información de la “Base de Datos de Conocimiento” suministrada por el cliente, que “es una base de datos que contiene todo el registro de errores conocidos”.[13]

- Diagnóstico y resolución en el primer contacto (First Call Resolution - FCR) de las incidencias relacionadas con el puesto de usuario, realizando toma de control remoto sobre la estación del usuario en caso de ser requerido.

- Documentar la información proporcionada por el usuario y las acciones tomadas para la solución en la herramienta de gestión.

- Escalar los incidentes o requerimientos a otros grupos resolutores según matriz de escalamiento definida entre el cliente e IBM.

- Solucionar la llamada cumpliendo los respectivos niveles de servicio acordados entre IBM y el cliente.

Una vez definidas las necesidades del cliente y responsabilidades de IBM, se debe proceder a evaluar el estatus actual de la mesa de servicio y basado en ello, proponer los planes de solución que tengan una buena factibilidad de implementación.

[13] Haren, Van. Operación del Servicio, Basada en ITIL®V3. Op. cit., p.132.

12. IDENTIFICACIÓN Y RECOLECCION DE INFORMACIÓN DE LOS FACTORES QUE EVALUAN EL DESEMPEÑO DE LA MESA DE AYUDA

"Dentro las funciones y procesos en la operación del servicio se debe incluir un grupo de trabajo encargo de la medición y control de servicios que se debe basar en un ciclo continuo de monitorización, elaboración de información e implementación de planes de acción para la optimización del servicio."[14]

Dentro de IBM existe un grupo de BackOffice que es el responsable de asegurar la integración de la Mesa de Servicio con el resto de los procesos y funciones del servicio, adicionalmente asegura la gestión de los niveles de servicio y la calidad de los procedimientos y actividades desarrollada por los agentes de la Mesa. Para evaluar la calidad del servicio se encarga de generar de informes en función de los datos obtenidos de las diferentes fuentes de entrada: herramientas de ticketing, herramientas de telefonía, herramientas WEB, entre otras, con el objetivo extraer los indicadores claves de rendimiento (KPI – Key Perfomance indicators), seguimiento de niveles de servicio contractuales (SLA – Service Level Agreements), así como la implementación de planes de acción de mejora continua.

Para lograr la medición de datos, el equipo de BackOffice utiliza dos fuentes de información, la primera es la base de datos donde se encuentran consolidados todos tickets creados por los agendes de la mesa de ayuda, en ésta se podrá encontrar detalladamente las indicaciones del usuario y las acciones realizadas para poder dar cierre a la llamada. La segunda fuente de información son los reportes generados por la central de telefonía especializada en la gestión y monitorización de Call Centers. Cada una de estas bases de datos servirá para extraer y modelar el sistema, con el fin de hallar soluciones basadas en la optimización del servicio y la rentabilidad de la empresa.

12.1. INFRAESTRUCTURA PARA LA ATENCIÓN Y MONITORIZACIÓN DE LLAMADAS.

IBM posee un distribuidor automático de llamadas (ACD - Automatic Call Distributor) que es el encargado de redirigir las llamadas entrantes a los empleados con mayor capacidad técnica dentro de la mesa de ayuda, monitorear la actividad en tiempo real y registrar la información en una base de datos.

Para este caso se utiliza una plata de distribución con el proveedor de servicios AVAYA que "Es una empresa de telecomunicaciones especializada en el sector de

[14] Haren, Van. Operación del Servicio, Basada en ITIL®V3. Op. cit., p.107.

la telefonía y centros de llamadas"[15]. La central ofrece paneles de monitoreo continuo que le dan una visión general al coordinador de la mesa de ayuda y de esta forma, se puedan ejecutar planes de mejora inmediatos que eviten el encolamiento dentro de las horas más críticas del día.

Figura 2. Representación de la herramienta de gestión del ACD de AVAYA y el teléfono 1400 Series Digital Deskphones usado en la mesa de ayuda de IBM.

Fuente: Imagen tomada de AVAYA. Avaya Aura® Call Center Elite [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.avaya.com/es/producto/avaya-aura-call-center-elite/

La monitorización continua de la mesa de ayuda es responsabilidad del coordinador, que se encarga de garantizar que la operación esté enfocada a suplir las necesidades del negocio, adicionalmente realiza la respetiva retroalimentación al personal. Sus funciones están orientadas a la satisfacción del usuario final y en el trabajo colaborativo con el coordinador del área de IT del cliente.

12.2. INFRAESTRUCUTRA TECNOLÓGICA SOFTWARE PARA LA CONSOLIDACIÓN Y CONTROL DE TICKETS.

Además del reporte de llamadas entrantes al sistema de la mesa de ayuda se debe evaluar la herramienta donde se consolida toda la información que es cargada por los agentes de la mesa de ayuda mientras atienden una llamada.

[15] AVAYA. Una reseña sobre nuestra compañía, Historia ITIL [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.avaya.com/es/about-avaya/our-company/company_overview/company-overview/

IBM incorpora la herramienta de Gestión del Servicio Smart Cloud Control Desk (Maximo) que es la única herramienta donde yace la información de los incidentes o requerimientos de los usuarios, incrementando así el valor del servicio.

Algunas de las características y funcionalidades de Máximo son las siguientes:

- Funciona a través de una plataforma WEB y proporciona acceso al módulo de auto servicio a usuarios finales, donde podrán acceder a soluciones de incidentes básicas.

- Identifica, clasifica y prioriza incidentes, permitiendo acceso rápido a una base de conocimiento que se alimenta en el proceso de gestión de incidentes.

- Utiliza Plantillas de Comunicaciones y Eventos para unificar criterios de contacto con usuarios.

- Permite llevar el registro completo de la jerarquía de Activos TI, incluyendo Equipos, Sistemas, Bases de Datos, Servicios y cualquier otro recurso, manteniendo su relación accesible desde cualquier punto de consulta.

- Provee capacidades de administración de incidentes y problemas y unifica los procesos claves de soporte de servicios

Esta herramienta permite aumentar la eficiencia en los servicios, reduce las interrupciones y perfecciona las operaciones de la mesa de servicio, siempre y cuando se sepa interpretar la información de forma adecuada para le ejecución de tareas complementarias.

13. TRATAMIENTO Y ANÁLISIS DE INFORMACIÓN

13.1. DISTRIBUIDOR AUTOMÁTICO DE LLAMADAS Y HERRAMIENTA DE GESTIÓN DE TICKETS

Una vez seleccionadas las fuentes de información con las que se va a trabajar, se procederá a analizar la información y para ello se deben utilizar herramientas capaces de interpretar gran cantidad de datos representándola de forma efectiva, es decir que se puedan visualizar los indicadores clave de rendimiento y así identificar fácilmente los posibles planes de acción para la implementación de la solución.

Para analizar el rendimiento de la mesa de ayuda es necesario tener en cuenta los siguientes indicadores de calidad:

Tasa de abandono: Se entiende por tasa de abandono, la relación porcentual entre la cantidad de llamadas no atendidas y el total de llamadas recibidas por la Mesa de servicio en un periodo de tiempo determinado, ver *Ecuación 1*. La tasa de abandono debe ser inferior o igual al 5%, ver *Ecuación2*. Para extraer este indicador es necesario contar con la totalidad de las llamadas entrantes a la mesa y las que fueron abandonadas pero su tiempo de espera fue superior a 15 segundos, ya que de lo contrario cualquier usuario pudiese llamar y colgar instantáneamente solo para afectar los indicadores de calidad.

$$tasa\ de\ abandono = \frac{llamadas\ abandonas}{llamadas\ entrantes\ a\ la\ mesa} \quad (1)$$

$$tasa\ de\ abandono \leq 5\% \quad (2)$$

Velocidad de Respuesta. Se entiende como Velocidad Respuesta al tiempo utilizado por la Mesa de Servicio para responder las llamadas, siendo el tiempo requerido máximo de 15 segundos para mínimo el 85% de las llamadas, ver *Ecuación 3 y 4.*

$$velocidad\ de\ respuesta = \frac{llamadas\ contestadas\ en\ menos\ de\ 15\ segundos}{llamadas\ entrantes\ a\ la\ mesa} \quad (3)$$

$$velocidad\ de\ respuesta \geq 85\% \quad (4)$$

Disponibilidad: Se entiende por disponibilidad de la Mesa de Servicio al porcentaje del tiempo operativo disponible de atención sobre el tiempo de soporte ofrecido. El nivel de disponibilidad de la Mesa de Servicio debe ser mínimo del 99,96% de disponibilidad, ver *Ecuación 5 y 6.*

$$disponibilidad = \frac{llamadas\ contestadas}{llamadas\ entrantes\ a\ la\ mesa} \quad (5)$$

$$disponibilidad \geq 99{,}96 \quad (6)$$

Una vez definida la información que se debe extraer de la base de datos, se procede a extraer un reporte de la herramienta obteniendo la información que se observa en la Tabla *1.*

Tabla 1. Base de datos extraída de la central de distribución de llamadas AVAYA.

Día	Calendario	Llamadas ofrecidas	Llamadas contestadas	Contestadas en <=15_segundos	Abandonas en >5_segundos
LUNES	04/01/2017	377	298	181	74
MARTES	05/01/2017	267	220	141	46
MIÉRCOLES	06/01/2017	198	186	168	11
JUEVES	07/01/2017	231	212	163	18
VIERNES	08/01/2017	236	218	171	18
MARTES	12/01/2017	446	291	134	150
MIÉRCOLES	13/01/2017	496	246	71	241
JUEVES	14/01/2017	307	225	106	77
VIERNES	15/01/2017	236	217	183	17
LUNES	18/01/2017	393	300	160	88
MARTES	19/01/2017	359	257	147	92
MIÉRCOLES	20/01/2017	347	220	96	127
JUEVES	21/01/2017	227	205	144	18
VIERNES	22/01/2017	237	211	148	26
LUNES	25/01/2017	421	300	137	119
MARTES	26/01/2017	275	235	147	39
MIÉRCOLES	27/01/2017	243	207	147	34
JUEVES	28/01/2017	234	208	151	25
VIERNES	29/01/2017	219	210	179	8
TOTAL	TOTAL	5749	4466	2774	1228

Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

Ya con la información necesaria se puede proceder a aplicar la Ecuaciones 1, Ecuación 3 y Ecuación 5 para verificar las métricas de calidad que evaluarían la mesa de ayuda, esta información se puede visualizar en la Tabla 2.

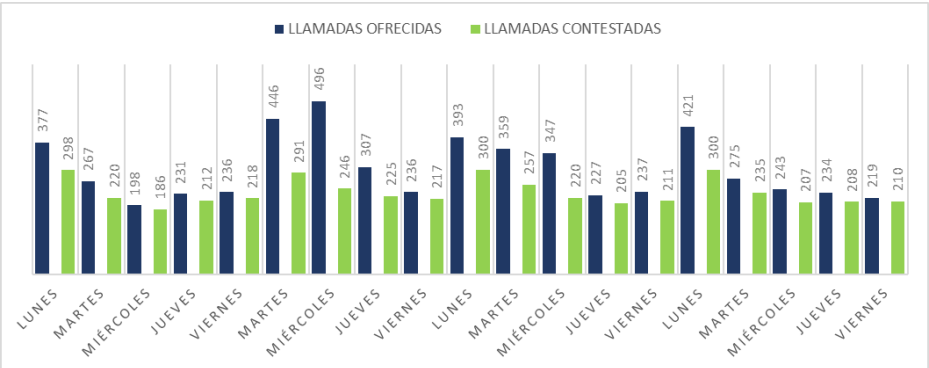
Tabla 2. Indicadores diarios de calidad de la mesa de ayuda.

Día	Calendario	Disponibilidad	Velocidad de respuesta	Tasa de abandono
LUNES	04/01/2017	79%	61%	20%
MARTES	05/01/2017	82%	64%	17%
MIÉRCOLES	06/01/2017	94%	90%	6%
JUEVES	07/01/2017	92%	77%	8%
VIERNES	08/01/2017	92%	78%	8%
MARTES	12/01/2017	65%	46%	34%
MIÉRCOLES	13/01/2017	50%	29%	49%
JUEVES	14/01/2017	73%	47%	25%
VIERNES	15/01/2017	92%	84%	7%
LUNES	18/01/2017	76%	53%	22%
MARTES	19/01/2017	72%	57%	26%
MIÉRCOLES	20/01/2017	63%	44%	37%
JUEVES	21/01/2017	90%	70%	8%
VIERNES	22/01/2017	89%	70%	11%
LUNES	25/01/2017	71%	46%	28%
MARTES	26/01/2017	85%	63%	14%
MIÉRCOLES	27/01/2017	85%	71%	14%
JUEVES	28/01/2017	89%	73%	11%
VIERNES	29/01/2017	96%	85%	4%
TOTAL	TOTAL	78%	62%	21%

Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

En la Figura 3 se puede observar que hay una tendencia donde los primero días de la semana se presenta una mayor cantidad de llamadas y así mismo comparando con la Tabla 3 se evidencia que la desviación con respecto a la media es de 70, 143, 90 y e incluso 118 llamadas, pero la capacidad de la mesa de ayuda no aumenta proporcionalmente ya que respectivamente se contestan 298, 291 y 300 llamadas, lo que lleva a pensar que se podría destinar un agente adicional el primer día de la semana para que soporte el equipo de la mesa de ayuda.

Figura 3. Diagrama de barras representando llamadas ofrecidas y contestadas vs días de la semana.



Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

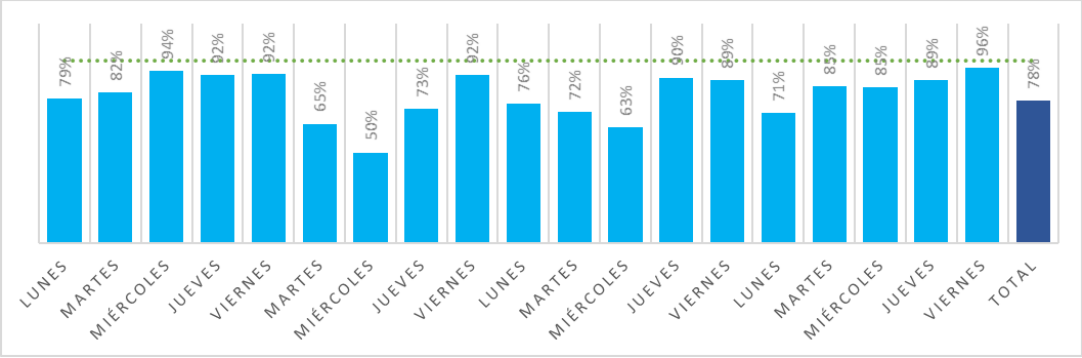
Tabla 3. Promedio y varianza de llamadas entrantes a la mesa de ayuda según métricas mensuales.

	Llamadas ofrecidas	Llamadas contestadas	Contestadas en <=15_segundos	Abandonas en >5_segundos
Total	5749.0	4466.0	2774.0	1228.0
Promedio diario	302.6	235.1	146.0	64.6
Varianza mensual	89.2	36.4	29.1	60.9

Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

En la Figura 5 se puede apreciar que el target de disponibilidad no se cumple en ningún día del mes, lo que da razón a implementar un plan de mejora o una herramienta de soporte adicional a los usuarios para que no colapsen diariamente la central telefónica. A pesar de que se incurrirían en costos adicionales, se podría estar evitando penalidades de incumplimiento, baja rentabilidad e insatisfacción del cliente.

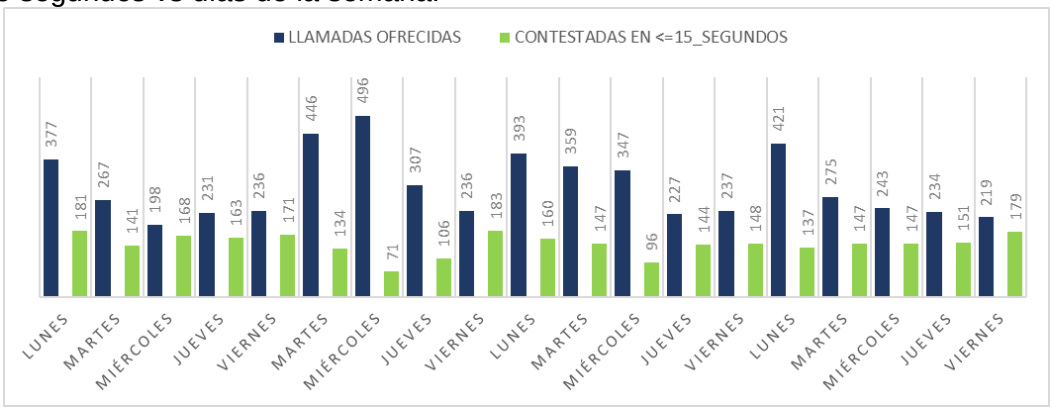
Figura 4. Diagrama de barras representando la tasa de disponibilidad con target señalado del 99.6%.



Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

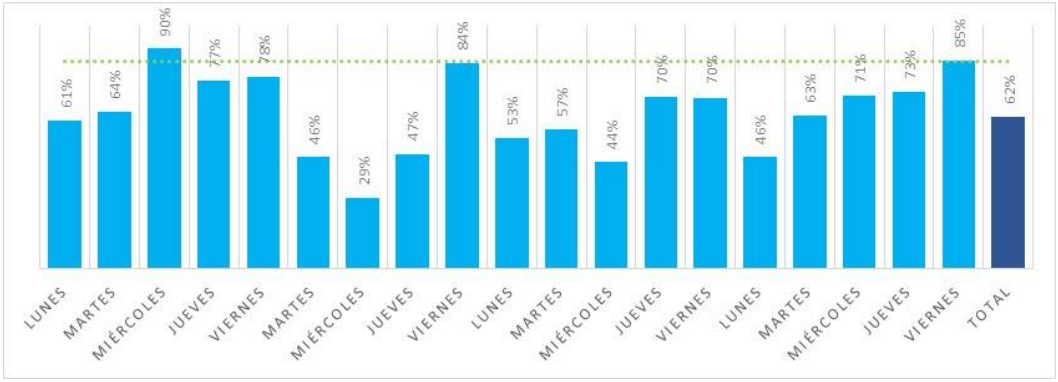
La tasa de velocidad de respuesta es un factor que puede ayudar a evidenciar el encolamiento que se presenta en la mesa de ayuda. Tal como se puede apreciar en la Figura 6 existe una capacidad limitada de contestar diariamente las llamadas cumpliendo el nivel de servicio acordado con el cliente que son 15 segundos máximo de espera, adicionalmente la Tabla 3 soporta la hipótesis de la capacidad limitada debido a que la varianza es tan solo de 30, la más pequeña entre los indicadores. Con estos datos se podría inferir que el tiempo de atención de cada llamada es similar y constante durante del día, debido a que la varianza de la cantidad de llamadas que cumplen la velocidad de respuesta estipulada por el cliente es mínima. Adicionalmente dado que la capacidad es limitada pero la cantidad de llamadas entrantes no lo es, tal como se observa en la Figura 7 la tasa de cumplimiento si varía constantemente con hasta 56 puntos de diferencia para cumplir el target del cliente.

Figura 5. Diagrama de barras representando llamadas ofrecidas y contestadas en menos de 15 segundos vs días de la semana.



Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

Figura 6. Diagrama de barras representando la tasa de velocidad de respuesta diaria con target señalado del 85%.



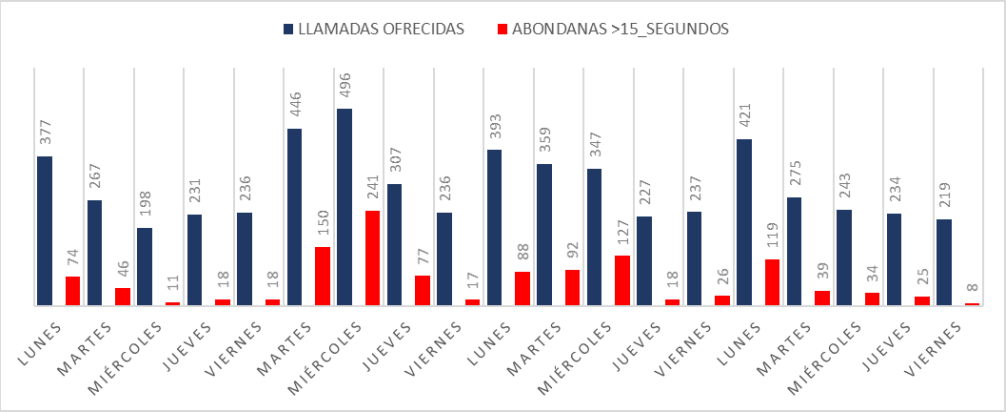
Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

Hay otro conjunto de llamadas que no solo ayudan a evidenciar el encolamiento sino la sobrecarga de llamadas entrantes en determinados periodos de tiempo que pueden ocasionar largos tiempos de espera y así mismo abandono, es decir los incidentes masivos, para este tipo de llamadas no hay ningún tipo de contingencia y son difícilmente controlables debido a que se pueden ocasionados por fallas en herramientas externas o la misma infraestructura de internet o redes de energía. La implementación de una solución que disminuya la carga de los agentes podría darles la capacidad de estar más preparados a la atención de este tipo de eventos ya que tal como se puede apreciar en la Figura 9 no se está cumpliendo con el target menor del 5% en ningún día de la semana.

El seguimiento mensual de la cantidad de llamadas abandonas es un soporte a la hora de justificar el incumplimiento en determinadas ocasiones, como por ejemplo el segundo Miércoles del mes donde se ve un abondo de 241 llamadas, Figura 9, tuvo la menor tasa de cumplimiento de velocidad de respuesta del mes que fue del 29%, Figura 7, de lo que se puede inferir que adicional a que se presentó un

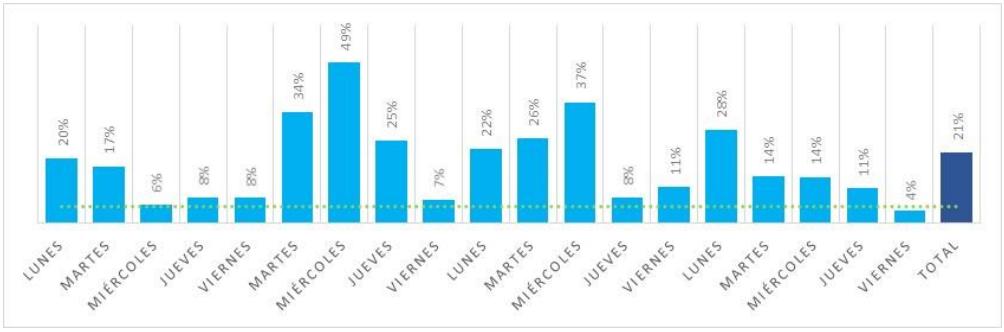
masivo, la solución de este inconveniente fue mucho más demorado del promedio de incidentes comunes y lo que igualmente genera la necesidad de crear un plan inmediato para el tipo de incidencia que se pudo haber presentado.

Figura 7. Diagrama de barras representando llamadas ofrecidas y no contestadas con tiempo de espera superior a 15 segundos vs días de la semana.



Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

Figura 8. Diagrama de barras representando la tasa de abandono diaria con el target señalado del 5%.



Fuente: Elaboración del autor, interpretación basada en ACD AVAYA.

Una vez analizados los datos del distribuidor automático de llamadas de AVAYA, donde se puede obtener una visión general de las llamadas entrantes, se procederá a tomar la segunda base de datos que es la generada por la herramienta de tickets. La información almacenada allí no solamente es por gestión de llamadas telefónicas contestadas, sino también por Chat, Email y Catálogo de autoservicio. En la Tabla 4 se aprecia que la mayor gestión la deben realizar los agentes que suplen los servicios del cliente vía telefónica, es decir que igualmente faltaría una campaña de concientización donde se informe que este tipo de medio solo puede ser utilizado cuando el usuario ha acabado todos los recursos por medio propio para solucionar la incidencia de prioridad baja ó media y podrá ser usado de forma directa con incidencias de prioridad 3. Las solicitudes que puedan tener tiempos de respuesta de una semana se deben gestionar por correo electrónico. Tal como se observa en la Figura 5. se alcanzaron a tener 134

llamadas por concepto de una consulta, práctica mal realizada por los usuarios ya que pueden estar disminuyendo la capacidad del servicio por una mala utilización que es causada por la desinformación.

Tabla 4. Cantidad de tickets ingresados a la mesa de ayuda según el tipo de ingreso.

FORMA DE INGRESO	TICKETS
CHAT	53
EMAIL	4315
PHONECALL	4569
SELFSERVICE	71
SERVICECATALOG	544
TOTAL	**9552**

Fuente: Elaboración del autor, interpretación basada en MAXIMO.

Tabla 5. Cantidad de tickets según el tipo de servicio atendido telefónicamente por los agentes de le mesa de ayuda.

TIPO DE SERVICIO	TICKETS
SOLICITUD	1700
INCIDENTE	2735
CONSULTA	134
TOTAL	**4569**

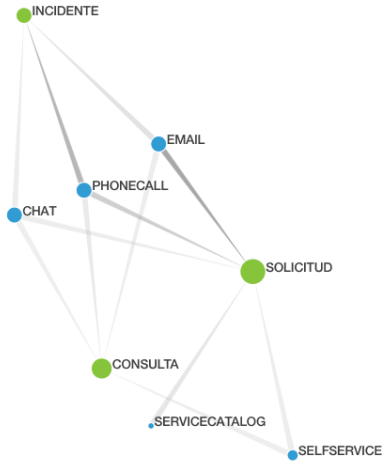
Fuente: Elaboración del autor, interpretación basada en MAXIMO.

Dado que para cada tipo de servicio se tuvieron todas las formas de ingreso posibles y para cada medio de contacto se creó de igual forma todo tipo de servicio, la mejor forma de analizar los datos en un tipo de gráfica relacional, con la de Watson Analytics; servicio inteligente de análisis y visualización de datos, capaz de descubrir rápidamente patrones y el significado de la información[16]. Para interpretar la Figura 9 se debe tener en cuenta que el radio de las esferas verdes es proporcional a la cantidad de parámetros relacionados y la intensidad del color de los enlaces es proporcional a la cantidad de tickets.

La primera interpretación que se realiza es que las solicitudes se están creando por todos los tipos de contacto posibles, pero las consultas no, situación para la cual se debe presentar el plan de mejora anteriormente mencionado donde el usuario no se comunique al centro de llamadas para solucionar dudas de dominio básico. A pesar de que hay gran cantidad de solicitudes gestionadas vía Email, también se evidencia una cantidad similar vía telefónica, que no se debería estar presentando por las condiciones de uso de la mesa anteriormente mencionadas.

[16] IBM Market Place. IBM Watson Analytics, What can I do for Business [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.ibm.com/ms-en/marketplace/watson-analytics

Figura 9. Relación entre el tipo de servicio y tipo de contacto en la mesa de ayuda.



Fuente: Elaboración del autor con la plataforma de análisis de información Watson Analytics.

El objetivo principal de una mesa de ayuda es lograr una centralización y solución de incidencias en el primer punto de contacto con una tasa de solución mayor a la del 90%, situación que no se está presentando como se puede observar en la Tabla 6. Esto se podría interpretar como una mala estructura en el área de IT debido a que un 35% de tickets se están escalando a otros grupos de trabajo que vendrían siendo fundamentales para la solución de la llamada pero entorpecen el flujo de trabajo en la mesa de ayuda, debido a que el desgaste operativo en el primer contacto con la documentación y servicio al cliente podría estar ocasionando encolamientos, como segundo plan de mejora se deberían incorporar los grupos resolutores con mayor alcance a la mesa de ayuda, así poder lograr una mejor centralización y distribución de carga laboral.

Tabla 6. Cantidad de tickets gestionados por grupo resolutor.

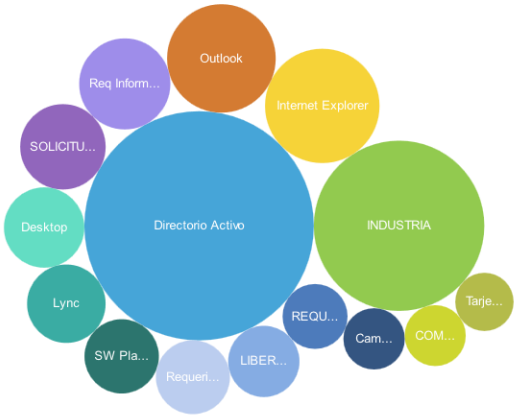
GRUPO RESOLUTOR	TICKETS	TICKETS %
MESA DE SERVICIO	2143	47%
OTROS	1580	35%
SOPORTE EN SITIO	846	19%
TOTAL	**4569**	**100%**

Fuente: Elaboración del autor, interpretación basada en MAXIMO.

En la Figura 10. Se evidencia que el tipo de plataforma por la cual se crean más llamadas es el directorio activo, donde básicamente la solución radica en el desbloqueo del usuario debido a múltiples intentos de inicio de sesión con contraseña incorrecta, en este punto también se podría evaluar un tercer plan de solución donde se disminuya la carga operativa del agente de ayuda, implementando una interfaz que agregue preguntas de seguridad para que el mismo usuario pueda reestablecer la contraseña o desbloquear su ID. Para la

solución de llamadas relacionadas al otro tipo de plataformas es donde se planea implementar un sistema de diagnóstico con interfaz al usuario donde se cargue la información de los incidentes más repetitivos y así mismo pueda acceder a ella, de forma más rápida.

Figura 10. Representación en burbuja relacionando el TOP 15 por cantidad de tickets de tipos de plataforma intervenidos por la mesa de ayuda.



Fuente: Elaboración del autor con la plataforma de análisis de información Watson Analytics.

Con la información de la Figura 11 se puede evidenciar el mal uso del centro de llamadas por parte del usuario, debido a que para incidencias de tipo 1 y 2 se está estableciendo contacto vía telefónica. Adicionalmente se aprecia que para las incidencias de tipo 3, hay cierta proporción donde no se cierre la llamada en el primer contacto, esto se infiere debido a la gran cantidad de escalamiento realizados. El alcance de los agentes también se está viendo limitado lo que refuerza la idea de incluir los demás grupos de solución al de la mesa de ayuda y de esa forma disminuir el tiempo de solución de un ticket para suplir las necesidades del cliente.

Figura 11. Representación de la cantidad de escalamientos vs prioridad según el tipo de contacto con la mesa de ayuda.

Fuente: Elaboración del autor con la plataforma de análisis de información Watson Analytics.

13.2. MODELAMIENTO DE LA MESA DE AYUDA Y SIMULACIÓN EN EL SOFTWARE ARENA

Para realizar la simulación en el software Arena es necesario contar con la tasa de entrada de las llamadas y la tasa de atención de cada de uno de los agentes. Para realizar el modelamiento de cada uno de los agentes se tomó la base de datos de MÁXIMO, es decir la herramienta de gestión tickets y se halló la cantidad de servicios atendidos por cada agente en el mes de muestra.

Tabla 7. Cantidad de tickets atendidos por los agentes de la mesa de ayuda en el mes de Enero.

AGENTE CREADOR	TICKETS ENERO
Augusto Orjuela	373
Camilo Fajardo	405
Christian Velez	413
Edith Palacio	349
Fabio Ochoa	350
Gian Ramirez	386
Hermes Pinto	342
Jose Martinez	420
July Pena	369
Lucas Martinez	435
Raul Coy	347
Ricardo Diaz	380
TOTAL	**4569**

Fuente: Elaboración del autor.

Dado que fue en enero de 2016 se debe tener en cuenta que la cantidad de días hábiles es de 19 y la duración de la jornada laboral es de 8 horas, con lo cual se calcula la tasa media de atención por hora de cada uno de los agentes, ver Tabla 9.

Para calcular la tasa de llegada de los usuarios, es necesario remitirse a la base de datos del distribuidor de llamadas AVAYA y teniendo en cuenta la información mencionada anteriormente, se calcula la tasa de entrada por hora, ver Tabla 8.

Tabla 8. Cantidad de llamadas entrantes a la mesa de ayuda, tasa media diaria y tasa media por hora

Descripción	TOTAL	MEDIA DIARIA	MEDIA POR HORA
Llamadas entrantes en el mes	5749	302.5789	37.8224

Fuente: Elaboración del autor

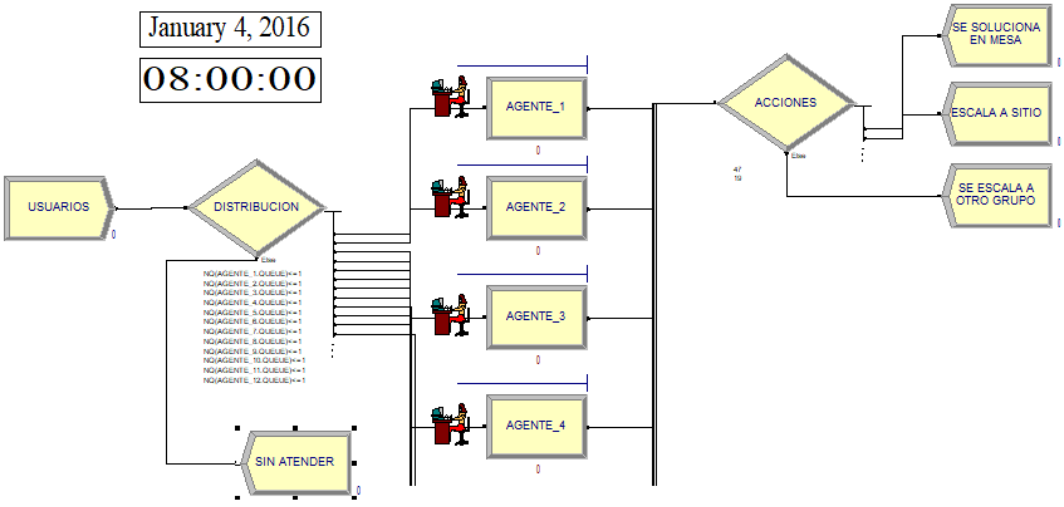
Tabla 9. Cantidad de tickets por día y media de tickets por hora.

AGENTE CREADOR	TICKETS PROMEDIO POR DÍA	TICKETS PROMEDIO HORA
Augusto Orjuela	19.6	2.4539
Camilo Fajardo	21.3	2.6645
Christian Velez	21.7	2.7171
Edith Palacio	18.4	2.2961
Fabio Ochoa	18.4	2.3026
Gian Ramirez	20.3	2.5395
Hermes Pinto	18.0	2.2500
Jose Martinez	22.1	2.7632
July Pena	19.4	2.4276
Lucas Martinez	22.9	2.8618
Raul Coy	18.3	2.2829
Ricardo Diaz	20.0	2.5000
MEDIA DE TICKETS	**20.0395**	**2.5049**

Fuente: Elaboración del autor.

Una vez hallados los valores principales se diseña el modelo de atención en el software Arena, ver Figura 12. Donde el bloque "*usuarios"* contiene la tasa de entrada de llamadas, el bloque *distribución* se encarga de administrar la cola, asignando a cada agente un usuario según la disponibilidad que tenga.

Figura 12. Simulación de la mesa de ayuda de IBM en software de simulación Arena.

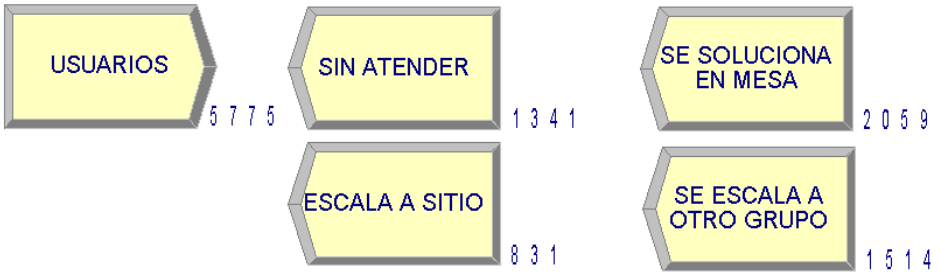


Fuente: Ambiente de simulación de software Arena.

En bloque "*distribución"* se encarga también de simular los usuarios que abandonan la llamada debido a que el sistema se encuentra encolado. Por última

instancia el bloque “acciones” se encarga de sacar la proporción de tickets que se serán resueltos por cada grupo resolutor, información que se configuró según Tabla 6.

Figura 13. Resultados del modelamiento de la mesa de ayuda de IBM con 12 agentes.



Fuente: Software de simulación Arena.

Una vez el software termina la simulación de los 19 días, se valida la cantidad de usuarios que terminaron en cada una de las estaciones, ver Figura 13 y se calcula la tasa de disponibilidad y la de abandono según la Ecuación 1 y la Ecuación 2. Se debe tener en cuenta que las llamadas solucionadas en mesa, escalas a sitio y demás grupos, fueron contestadas. Una vez obtenidos los indicadores de calidad de la mesa de ayuda simulada (ver Tabla 10), se validó que el modelo de simulación se comportó de forma semejante a la realidad, En la Tabla 2, se observa que la tasa de disponibilidad y de abondo fue de 78% y 21% respectivamente; es decir tan solo dos puntos de diferencia con el modelo real.

Tabla 10. Comparación de los indicadores de calidad de la mesa de ayuda, basado en la simulación Arena.

Entrantes	Contestadas	Sin Atender	Tasa de Disponibilidad		Tasa de Abandono	
			Simulación	Target	Simulación	Target
5775	4404	1341	***76%***	99.96%	***23%***	5.00%

Fuente: Elaboración del autor.

Con el comportamiento del modelo del sistema muy cercano a la realidad, se podrán probar los diferentes planes de solución, observar su respuesta y ejecutar acciones basadas en resultados más fiables.

14. PRESENTACIÓN DEL PLAN DE MEJORA

Para la implementación de un plan de solución se tendrán en cuenta dos factores; la optimización de los costos, los cuales se van a obtener mediante el análisis por medio de la teoría de colas y la reducción de encolamiento que se mitigará mediante la implementación de un sistema de diagnóstica basado en el entrenamiento de una red neuronal y un chatbot que se programará para que converse con los usuarios, a fin de solventar sus necesidades en el área de IT.

14.1. SOLUCIÓN DE PROBLEMAS DE ENCOLAMIENTO MEDIANTE LA TEORÍA DE COLAS

Proporcionar demasiada capacidad de servicio para operar el sistema implica costos excesivos; pero si no se cuenta con suficiente capacidad de servicio surgen esperas excesivas con todas sus desafortunadas consecuencias. Los modelos permiten encontrar un balance adecuado entre el costo de servicio y el costo de espera.[17]

Un sistema de colas cuenta con el siguiente flujo de operación: “Los clientes que requieren un servicio se generan en el tiempo en una fuente de entrada, luego, entran al sistema y se unen a una cola. En determinado momento se selecciona un miembro de la cola para proporcionarle el servicio mediante alguna regla conocida como disciplina de la cola. Se lleva a cabo el servicio que el cliente requiere mediante un mecanismo de servicio, y después el cliente sale del sistema de colas.[18]

Para la mesa de ayuda en IBM se tiene una disciplina de cola donde el primero en llegar es el primero en atender es decir que no existe ninguna clase de priorización, adicionalmente el mecanismo de servicio son los agentes contestando las llamadas.

Se pueden usar modelos de líneas de espera para equilibrar las ganancias que podrían lograrse al aumentar la eficiencia del sistema contra los costos de hacerlo. Además, se debe considerar los costos de no hacer mejoras al sistema: Las largas líneas de espera o los largos tiempos de espera pueden hacer que los clientes abandonen el sistema o se nieguen a utilizarlo. Por ende, cada una de las características de un sistema de colas puede ser analizada e interpretada según corresponda:

[17] HILLIER, Frederick S; LIEBERMAN, Gerald J. Introducción a la investigación de operaciones. Op. cit., P.708.
[18] Ibíd., p.709.

- Longitud de línea: El número de clientes en la línea de espera puede interpretarse según las características del servicio. Las líneas cortas pueden significar un buen servicio al cliente o demasiada capacidad. De manera similar, las líneas largas podrían indicar una baja eficiencia del servidor o la necesidad de aumentar la capacidad.

- Número de clientes en el sistema. El número de clientes en línea y en servicio también se relaciona con la eficiencia y capacidad del servicio. Una gran cantidad de clientes en el sistema provoca congestión y puede generar insatisfacción del cliente, a menos que se agregue más capacidad.

- Tiempo de espera en línea. Las líneas largas no siempre significan largos tiempos de espera. Si la tasa de servicio es rápida, una larga cola puede servir de manera eficiente. Sin embargo, cuando el tiempo de espera parece largo, los clientes perciben que la calidad del servicio es pobre. Entonces se puede tratar de cambiar la tasa de llegada de clientes o diseñar el sistema para que los tiempos de espera largos parezcan más cortos de lo que realmente son.

- Tiempo total en el sistema. El tiempo total transcurrido desde la entrada al sistema hasta la salida del sistema puede indicar problemas con los clientes, la eficiencia del servidor o la capacidad. Si algunos clientes pasan demasiado tiempo en el sistema de servicio, puede ser necesario cambiar la disciplina de prioridad, aumentar la productividad o ajustar la capacidad de alguna manera.

- Utilización de la infraestructura del servicio. La utilización colectiva de las instalaciones de servicio refleja el porcentaje de tiempo que están ocupados. El objetivo de la administración es mantener una alta utilización y rentabilidad sin afectar adversamente las otras características operativas.

El mejor método para analizar un problema de línea de espera es relacionar las cinco características operativas y su costo. Sin embargo, definir un costo exacto en ciertas características (como el tiempo de espera de un comprador en un supermercado) es difícil. En tales casos, un analista administrativo debe considerar el costo de implementar la alternativa, bajo consideración del costo en caso de no realizar el cambio.[19]

Para que el modelamiento sea útil, debe ser lo suficientemente realista como para que el modelo proporcione predicciones razonables, pero al mismo tiempo debe ser lo suficientemente sencillo para que sea matemáticamente manejable. El modelo M/M/s supone que todos los tiempos entre llegadas son independientes e idénticamente distribuidos de acuerdo con una distribución exponencial (es decir, proceso Poisson, cuyas llegadas al sistema ocurren de manera aleatoria, pero con

[19] KRAJEWSKI, Lee. Operations Management. 10 ed. Kendallville: Pearson Education Limited, 2013. p. 250.

cierta tasa media fija y sin que importe cuántos clientes están ya ahí) y que el número de servidores es s (cualquier entero positivo).[20]

Una vez entendidos los conceptos, se aplicarán los cálculos que ayudarán a hallar los valores más significativos del sistema:

$$P_0 = \frac{1}{1 + \sum_{n=0}^{s-1} \frac{(\lambda/\mu)^n}{n!} + \frac{(\lambda/\mu)^s}{s!} * \frac{1}{1 - \lambda/s\mu}} \quad (7)[21]$$

$$Lq = \frac{P_0(\lambda/\mu)^s \rho}{s!\,(1-\rho)^2} \quad (8)[22]$$

$$L = L_q + \frac{\lambda}{\mu} \quad (9)[23]$$

donde

λ: Tasa media de llegada.
μ: Tasa media de servicio en todo el sistema
s: Número de servidores.
ρ: Factor de utilización $\lambda/s\mu$
P_0: Probabilidad de que haya 0 clientes en el sistema.
L_q: Longitud esperada de la cola
L: Número esperado de clientes en el sistema
$E(CT)$: Costo total esperado por unidad de tiempo,
$E(CS)$: Costo de servicio esperado por unidad de tiempo,
$E(CW)$: Costo de espera por unidad de tiempo.

El objetivo es seleccionar el número de servidores para minimizar el costo total esperado por unidad de tiempo:

$$E(CT) = E(CS)\ E(CW) \quad (10)$$

[20] HILLIER, Frederick S; LIEBERMAN, Gerald J. Introducción a la investigación de operaciones. Op. cit., p.725.
[21] Ibíd., p729.
[22] Ibíd., p.730.
[23] Ibíd., p.730.

Cuando el costo de cada uno de los servidores es el mismo, el costo de servicio es:

$$E(CS) = C_s s \qquad (11)$$

Por lo tanto, cuando el costo de espera es proporcional a la cantidad de clientes en el sistema, este costo puede expresarse como:

$$E(CW) = C_w L \qquad (12)$$

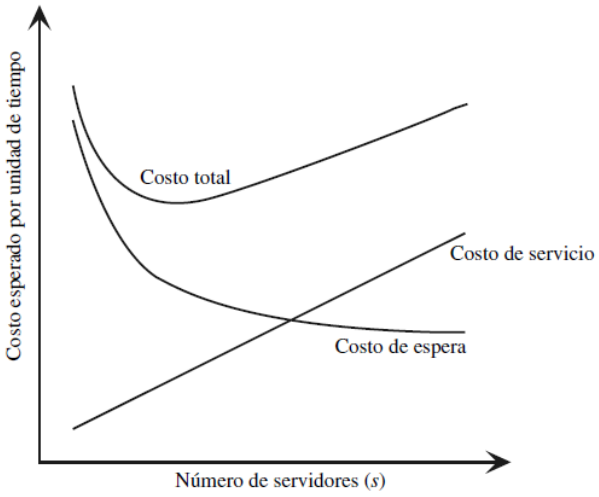
donde C_w es el costo de espera por unidad de tiempo de cada cliente en el sistema de colas.

En consecuencia, después de estimar las constantes, C_s y C_w, la meta es elegir el valor de s para minimizar:

$$E(CT) = C_s s + C_w L \qquad (13)$$

En la Figura 14 se puede apreciar el comportamiento de los costos del servicio y de la espera, donde se debe hallar el punto mínimo para la optimización del sistema.

Figura 14. Comportamiento del costo total según la variación del número de servidores.

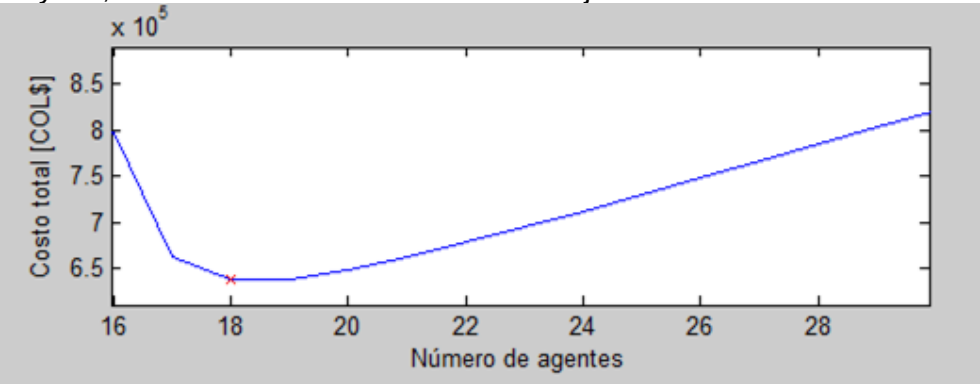


Fuente: Tomado de HILLIER, Frederick S; LIEBERMAN, Gerald J. Introducción a la investigación de operaciones. 9 ed. Mexico DF: McGraw-Hill/Interamericana editores, 2010. p.755.

Para el caso en específico se discriminaron costos de operación que no dependen de la cantidad de servidores, como por ejemplo, el costo por hora del Team leader, coordinador o especialista, o bien sea el costo de infraestructura diferido por hora como lo es el de la WAN utilizada para el acceso de control remoto, IBM EndPoint Manager utilizado para la interconexión IP de las plantas telefónicas entre IBM y el cliente, el distribuidor automático de llamadas de AVAYA, los servidores dedicados al registro de eventos e información de las llamadas o la planta eléctrica de emergencia UPS.

La tasa de entrada de llamadas se extrae de la Tabla 8 donde ya se habían calculado los valores y la tasa de atención media del sistema de la Tabla 9. El costo del agente por hora se estableció en COP$18,000.00 y el costo del usuario esperando se definió según el costo de trabajo del usuario es decir el salario, pero en una compañía el rango salarial puede variar desde los COP$18,000.00 para un analista o hasta de COP$230,000.00 para un ejecutivo. Para cada caso en específico se calculó el costo total de la operación por hora y se obtuvo la variación financiera según la cantidad de servidores, ver Tabla 16 y Tabla 17, donde se halló que para una mesa de ayuda con atención a usuarios de menor promedio salarial se debe contratar a 18 agentes y los costos totales se verían reducidos a COP$638,090.00 por hora e incluso COP$96,989,680.00 por mes logrando así un balance entre el costo de operación y el costo de nómina de los usuarios que tuvieron afectación en su estación de trabajo.

Figura 15. Variación Costo total vs cantidad de servidores, en la situación actual de la mesa de ayuda, basada en el costo mínimo de trabajo de los usuarios.

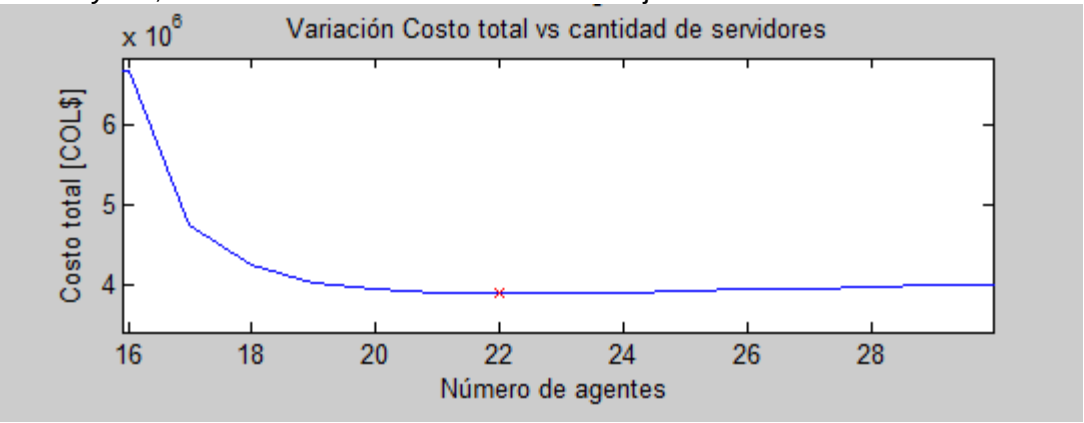


Fuente: Elaboración del autor.

Para una mesa de ayuda con atención a usuarios de alto promedio salarial se debe contratar a 22 agentes y los costos totales se verían reducidos a COP$3,907,200.00 por hora y COP$593,894,400.00 por mes, ver Figura 17.

Este plan de solución no es el definitivo ya que posterior a la implementación del plan mejora se podrían mitigar más incidencias, lo que disminuiría la cantidad de llamadas entrantes y así mismo se podría reducir la cantidad de agentes.

Figura 16.Variación Costo total vs cantidad de servidores, en la situación actual de la mesa de ayuda, basada en el costo máximo de trabajo de los usuarios.



Fuente: Elaboración del autor.

14.2. DISEÑO E IMPLEMENTACIÓN DEL SOFTWARE DE DIAGNÓSTICO BASADO EN EL ENTRENAMIENTO DE REDES NEURONALES

La tecnología neuronal trata de reproducir el proceso de solución de problemas del cerebro. Así como los humanos aplican el conocimiento ganado con la experiencia a nuevos problemas o situaciones, una red neuronal toma como ejemplos problemas resueltos para construir un sistema que toma decisiones y realiza clasificaciones. Los problemas adecuados para la solución neuronal son aquellos que no tienen solución computacional precisa o que requieren algoritmos muy extensos como en el caso del reconocimiento de imágenes.[24]

Una red neuronal con perceptrón multicapa es uno de los primeros esfuerzos por representar el aprendizaje supervisado, donde se emplean funciones de activación elementales de forma binaria en la mayoría de las ocasiones. Este elemento es un clasificador lineal con dos entradas y un punto de ajuste, siendo la salida el clasificador. La retropropagación (backpropagation) del error es un entrenamiento supervisado que se emplea para redes multicapa, donde se ajusta el valor de los pesos en función del error generado. Esta técnica es muy empleada ya que permite tener un método de optimización que se encuentra al definir el gradiente del error y minimizarlo con respecto a los parámetros de la red neuronal. [25]

Para la red neuronal de diagnóstico de problemas se usará una red especializada en el reconocimiento de patrones en Matlab, que son redes avanzadas que

[24] PONCE CRUZ, Pedro. Inteligencia Artificial con aplicaciones a la ingeniería. 1 ed. México D.F: Alfaomega Grupo Editor, 2010. p.6.
[25] Ibid. p.6.

pueden entrenarse para clasificar las entradas de acuerdo a diferentes agrupaciones de elementos.[26]

El proceso de clasificación parte del conocimiento de la existencia de un conjunto de clases, para determinar el conjunto de reglas, que permita asignar, a cada nueva observación la clase a la que pertenece. En este caso el tipo de aprendizaje a partir de ejemplos que ha de utilizarse es el aprendizaje supervisado, ya que se dispone de la información sobre la clase para cada ejemplo.[27]

El problema de valorar la calidad de un clasificador de tipo supervisado es que es necesario saber a priori el número de prototipos que se van a utilizar en la solución. En los sistemas de neuronas, es necesario saber cuántas agrupaciones existen, ya que hay que incluir una célula por cada agrupación. Como se desconoce totalmente la estructura de los datos, es necesario utilizar el método de prueba y error. Se han de realizar experimentos con diferente número de prototipos. Es aquí donde surge la dificultad de valorar los resultados. Por regla general, el número de errores disminuye al aumentar el número de prototipos, pero el poder de generalización del sistema disminuye. Si se tiene un sistema clasificador con pocos prototipos, pero un error alto, y otro con menor error, pero con muchos prototipos, la comparación entre ambos, en cuanto a calidad, es difícil de realizar, pues no existe ningún criterio objetivo para decidir que uno es mejor que otro.[28]

Tabla 11. Cantidad de tickets resueltos por la mesa de ayuda y el soporte en sitio según el tipo sistema operativo.

Grupo resolutor	Windows 8	Equipo Completo	Windows	Windows 7
I-IBM-SOPORTE_SITIO	43	34	12	12
I-IBM-MESA_DE_SERVICIO	43	4	22	20

Fuente: Elaboración del autor.

Para la definir la cantidad de agrupaciones en la red neuronal, es necesario identificar un grupo de incidencias que tenga una serie de síntomas similares, pero con soluciones diferentes y así crear una red que sea capaz de entrenarse y retroalimentarse según el ambiente de implementación. Para este caso, se concluyó que los problemas relacionados al sistema operativo tienen dichas características mencionadas anteriormente, por lo que fue necesario acudir a la herramienta de gestión de tickets (Máximo) e identificar la cantidad de problemas asociados, ver Tabla 11, donde 190 tickets se podrían mitigar con la implementación del software de diagnóstico.

[26] MATHWORKS. Patternnet, documentation [en linea], [revisado 2 de Octubre de 2017]. Disponible en Internet: https://es.mathworks.com/help/nnet/ref/patternnet.html

[27] ISASI VIÑUELA, Pedro. Redes Neuronales Artificiales, un Enfoque Práctico. 1 ed. Madrid: Pearson Educación S.A, 2004. p.199.

[28] ISASI VIÑUELA, Pedro. Redes Neuronales Artificiales, un Enfoque Práctico. 1 ed. Madrid: Pearson Educación S.A, 2004.p.200.

Para dicha población de casos se validó el método de solución, consignado en la Tabla 12. Posteriormente se procedió a configurar una serie de vectores de entrada y de salida de la red neuronal que corresponderían a los síntomas y a los métodos de solución.

Para el vector de salidas se utilizaron en total 10 posibles diagnósticos de solución, por ende, se necesitarían al menos 10 perceptrones, al realizar varias pruebas de rendimiento se valida que, con 30 perceptrones en una sola capa, bastan para entrenar el sistema y obtener los resultados deseados. Adicionalmente se utilizó el método de entrenamiento de Levenberg-Marquardt que es el algoritmo de retropropagación más rápido que tiene el "Neural Network Toolbox" de Matlab.[29] Se denomina entrenamiento al proceso de configuración de una red neuronal para que las entradas produzcan las salidas deseadas a través del fortalecimiento de las conexiones. Una forma de llevar esto a cabo es a partir del establecimiento de pesos conocidos con anterioridad, y otro método implica el uso de técnicas de retroalimentación y patrones de aprendizaje que cambian los pesos hasta encontrar los adecuados.[30]

Tabla 12. Cantidad de tickets según el método de solución relacionado a problemas del sistema operativo.

Solución	Tickets
Configurar Impresora	45
Mapear Unidad De Red	34
Reinicio de equipo	17
Error En Sistema Operativo	17
Ampliación de memoria ram	23
Actualizar Politicas	13
Ingresar Al Dominio	10
Limpieza de Temporales	7
Formateo de Equipo	6
Limpieza de CPU	6
Desfragmentar disco	5
Ejectutar chequeo antivirus	4
Cambio de disco duro	4
Grand Total	**190**

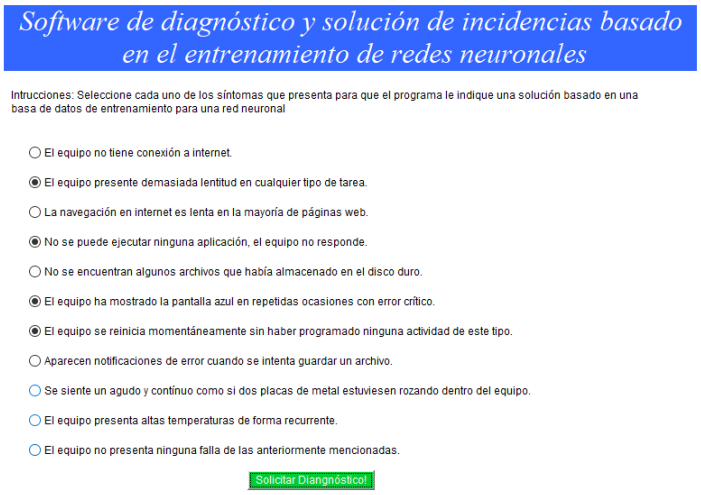
Fuente: Elaboración del autor.

El diseño de la interfaz gráfica del software de diagnóstico (Figura 17), cuenta un cheklist para que el usuario elija todos los síntomas que está presentando y que bien podría tambien ser consultado por personal técnico. Una vez el usuario da click sobre la opción "Solicitar diagnóstico", la red previamente entrenada arroja el resultado de lo que sería el diagnóstico, ver Figura 18.

[29] MATHWORKS, Levenberg-Marquardt backpropagation [en línea], [revisado el 2 de Octubre]. Encontrado en internet: https://es.mathworks.com/help/nnet/ref/trainlm.html

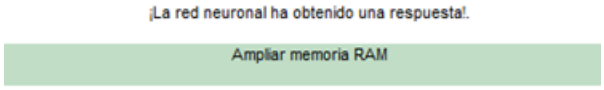
[30] PONCE CRUZ, Pedro. Inteligencia Artificial con aplicaciones a la ingeniería. 1 ed. México D.F: Alfaomega Grupo Editor, 2010. p.203.

Figura 17. Interfaz gráfica de software de diagnóstico de problemas basado en el entrenamiento de redes neuronales.



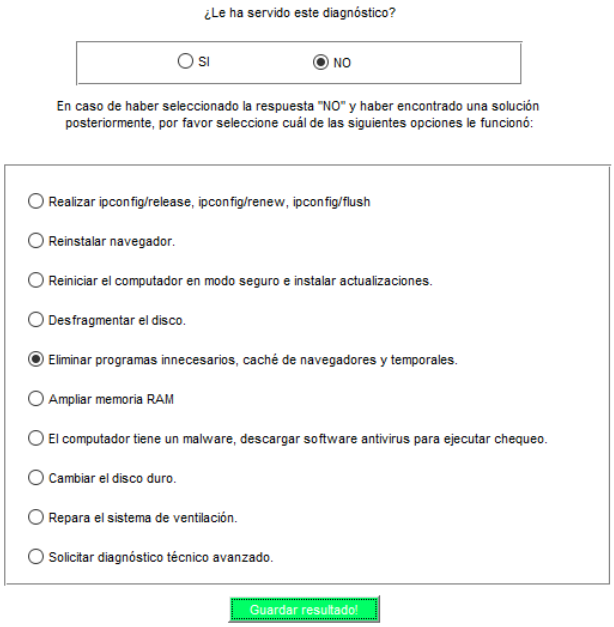
Fuente:Elaboración del autor.

Figura 18. Diagnóstico final de software basado en entrenamiento de redes neuronales.



Fuente: Elaboración del autor.

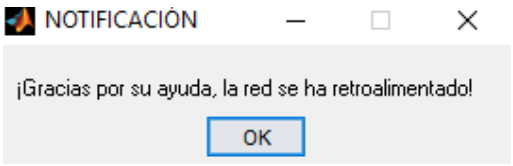
Figura 19. Mensaje de validación del software de diagnóstico de problemas para el reentrenamiento de la red neuronal



Fuente: Elaboración del autor.

Dado caso el usuario realice el diagnóstico que arrojó el software y no le funcione, esté puede seleccionar la opción que le funcionó, la red se volvería a entrenar y arrojaría una notificación al usuario, ver Figura 20. Dado que los equipos de los usuarios tienen las mismas características, el diagnóstico que se almacenó podría servirle a otra persona que presente los mismos síntomas.

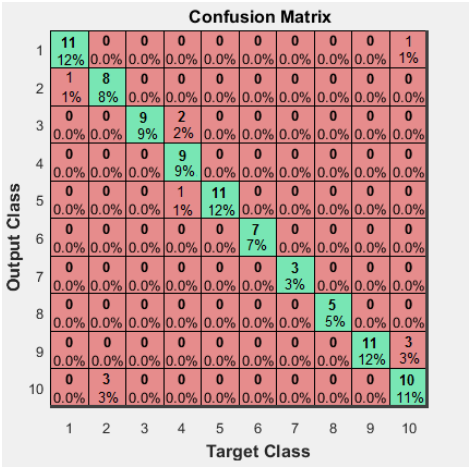
Figura 20. Ventana emergente de notificación de software indicando que una nueva red neuronal se ha entrenado



Fuente: Elaboración del autor.

Para validar el rendimiento de la red neuronal es necesario evaluar la matriz de confusión que se puede ver a continuación.

Figura 21. Matriz de confusión de la red neuronal.



Fuente: Toolbox Neural Networks de Matlab.

Donde basados en la teoría se extraerá la cantidad de verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos, con el objetivo de hallar la precisión, sensibilidad y especificidad de cada una de las clases de la red neuronal. Se puede evidenciar en la Tabla 13 que para las clases 4, 5, 7 y 8 existe una mayor tasa de sensibilidad lo que quiere decir que cuando el sistema indica que se debe realizar dicho diagnóstico hay una alta probabilidad de que efectivamente al usuario le funcione y no tenga la necesidad de entrar al sistema de espera. Para dicha de matriz de confusión se tiene una exactitud general del 88.42% y basados en la Tabla 12 donde se indica que 190 usuarios consultaron este tipo de problemas, se podría concluir que 167 llamadas mensuales hubiesen podido ser atendidas antes de entrar al sistema.

Tabla 13. Precisión, sensibilidad y especificidad de cada una de las clases de la red neuronal.

CLASE	FN	FP	TN	TP	Precisión	Sensibilidad	Especificidad
1	1	1	82	11	91.67%	91.67%	98.80%
2	1	3	83	8	72.73%	88.89%	96.51%
3	2	0	84	9	100.00%	81.82%	100.00%
4	0	2	84	9	81.82%	100.00%	97.67%
5	0	1	83	11	91.67%	100.00%	98.81%
6	1	0	87	7	100.00%	87.50%	100.00%
7	0	0	92	3	100.00%	100.00%	100.00%
8	0	0	90	5	100.00%	100.00%	100.00%
9	3	0	81	11	100.00%	78.57%	100.00%
10	3	4	78	10	71.43%	76.92%	95.12%

Fuente: Elaboración del autor.

14.3. DISEÑO E IMPLEMENTACIÓN DE CHATBOT PARA LA REDUCCIÓN DEL ENCOLAMAMIENTO EN LA MESA DE AYUDA.

Para la implementación del chatbot (chat computadorizado) se utilizó Bluemix (entorno de plataforma como servicio desarrollado por IBM con metodología de desarrollo DevOps de forma integrada para crear, ejecutar, desplegar y gestionar aplicaciones en la nube[31]), donde se utilizó la API (conjunto de funciones o métodos que pueden ser utilizados por otro software como una capa de abstracción[32]) de Watson Conversation que es un servicio de conversación donde se combina el aprendizaje automático, la comprensión del lenguaje natural y las herramientas de diálogo, integradas para crear flujos de conversación entre aplicaciones y sus usuarios.[33]

Para diseñar el flujo de conversación primero se debe pensar en las necesidades de los usuarios, para ello se utilizó la base de datos de gestión de llamadas, donde se identificaron las plataformas por las cuales se realizaban más consultas a la mesa de ayuda.

La elaboración de instrucciones para cada uno de las solicitudes o problemas que presentan los usuarios no es posible, debido a que en ciertos ámbitos se requiere de herramientas especializadas, permisos adicionales o aprobaciones previas, como lo es en el caso del Directorio Activo (peticiones para desbloquear usuario, reiniciar contraseña, entre otros), pero en otro tipo de plataformas como lo es

[31] IBM Cloud Platform. IBM Bluemix [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.ibm.com/cl-es/marketplace/cloud-platform

[32] RUDRAKSHI, Charu. API-fication, Core Building Block of the Digital Enterprise [en linea], Agosto de 2014 [revisado 2 de Octubre de 2017]. Disponible en Internet: http://www.hcltech.com/sites/default/files/apis_for_dsi.pdf

[33] IBM Developer Cloud. API Reference, Watson Conversation, Introduction [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.ibm.com/watson/developercloud/conversation/api/v1/?cm_mc_uid=30613969409014902027442&cm_mc_sid_50200000=1501388877&cm_mc_sid_52640000=1501388877#introduction

Outlook, Internet Explorer, Lync o la VPN, solo se requiere seguir cierto tipo de instrucciones para poder solucionar un incidente o lograr la utilización de una funcionalidad.

Tabla 13.Cantidad de tickets según el Top de plataformas intervenidas por la mesa de ayuda y sitio.

PLATAFORMA	TICKETS
Directorio Activo	632
Internet Explorer	295
Outlook	261
Req Informacion	214
Requerimiento no Estandar	98
Patch Cord	89
Windows 8	85
Camaras	85
Lync	68
Desktop	60
Navegadores	57
VPN UAG	57
Tarjeta Red	56
Bandeja	46
Inc Sw Negocio	44
SW Plataforma	42
Laptop	41
REQUERIMIENTO DE INFORMACION	39
EQUIPO COMPLETO	38
Google Chrome	34
Otros	648
TOTAL	**2989**

Fuente: Elaboración del autor.

Para obtener ideas acerca de la información con la cual se iba a construir la base de conocimiento del chatbot, se extrajeron los métodos de solución para cada una de las plataformas mencionadas anteriormente, ver Tabla 14, Tabla 15 y Tabla 16.

Una vez identificada la información se procedió a crear un diagrama de flujo con lo que sería una conversación entre el chatbot y el usuario, cabe resaltar que se realizó una versión de prueba (beta), debido a que la creación de un programa que contuviese una base de conocimiento completa, requiere de un equipo de trabajo especializado y programación durante un periodo de tiempo prolongado.

Tabla 14. Cantidad de tickets según el método de solución empleado en la plataforma Internet Explorer.

Internet Explorer	TICKETS
Reinstalar-Reparar	199
Realizar Mantenimiento	76
Internet Explorer	8
Instalar Complemento	7
Actualizar	2
Indicar Manejo	1
Desinstalar	1
Parametrizar/Configurar	1
TOTAL	**295**

Fuente: Elaboración del autor.

Tabla 15. Cantidad de tickets según el método de solución empleado en la plataforma Outlook.

Outlook	TICKETS
Configurar Aplicativo	141
Outlook	42
Error De Envio De Correo	23
Reiniciar Aplicacion	17
Configurar	14
No Ingresa Correo	5
Reinstalar-Reparar	4
Error En Archivado Local PST	4
Configurar archivo PST	3
Instalar	3
No Llegan Los Correos A Los Destinatarios	2
Indicar Manejo	2
Buzon Lleno	1
TOTAL	**261**

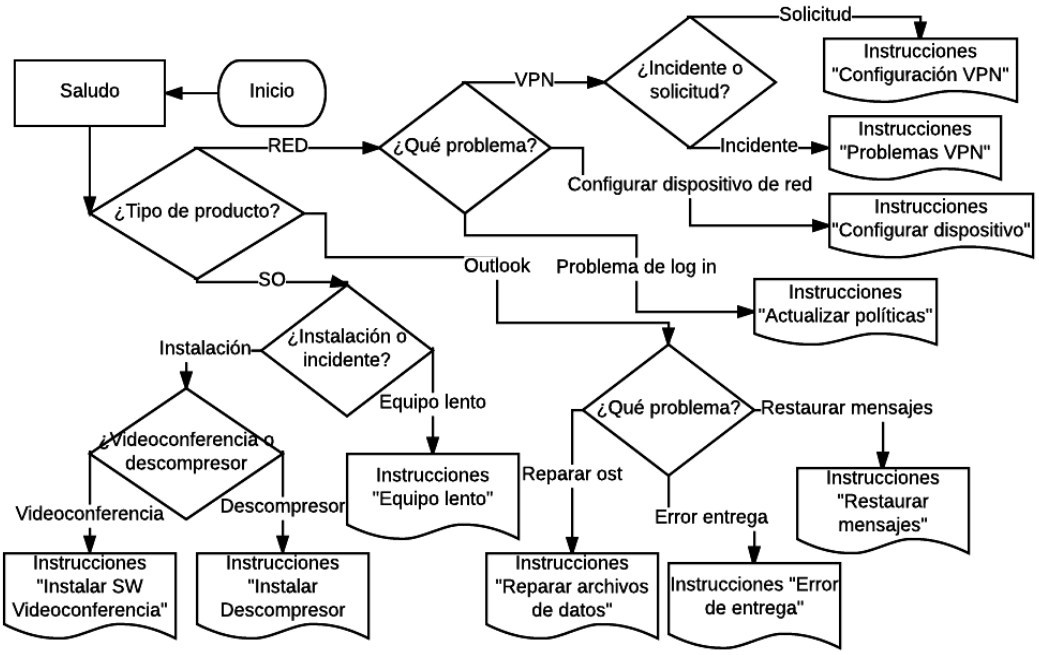
Fuente: Elaboración del autor.

Tabla 16. Cantidad de tickets según el método de solución empleado en la plataforma Lync.

Lync	TICKETS
Lync	25
Reinstalar-Reparar	15
Desbloqueo de Usuario	10
Ajustar configuracion	10
Cambio de Contrasena	4
Actualizar	2
Configurar	2
TOTAL	**68**

Fuente: Elaboración del autor.

Figura 21. Diagrama de flujo de chatbot implementando Watson Conversation.



Fuente: Elaboración del autor.

Una vez realizado el flujo de conversación, se debe entrenar el chatbot y para ello se requiere de ciertos parámetros de programación, que son los siguientes:

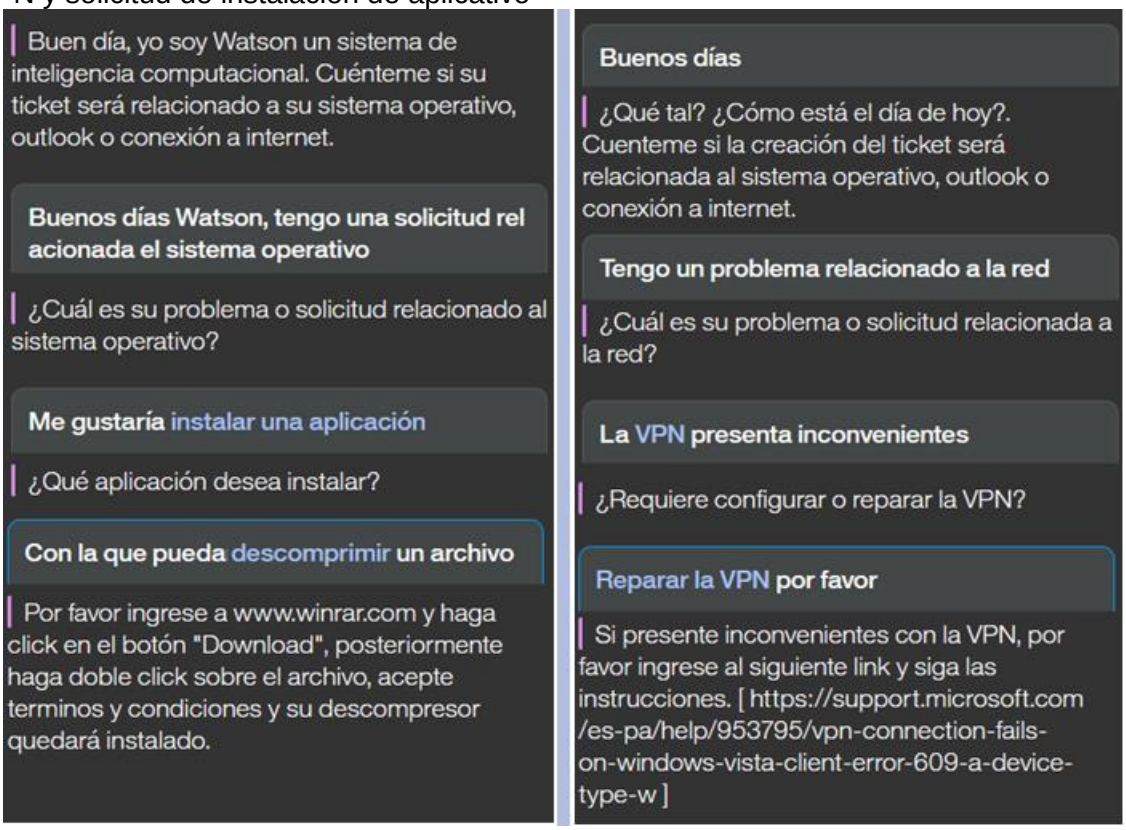
- Intents: Son las intenciones de los usuarios que sirven para diseñar respuestas automáticas a dichas consultas[34], con este tipo de parámetros usualmente se subdividen flujos de conversación. Por ejemplo, si el usuario pregunta ¿Cómo puedo instalar Lync? ó ¿Cómo puedo configurar la VPN?, el flujo de conversación cambiaría para cada uno de los contextos.

- Entities: una entidad representa un término u objeto que proporciona un contexto para una intención.[35] Por ejemplo, Outlook presenta Error 123456 ó Outlook presenta Error 64321 que ayudaría a identificar qué tipo de diagnóstico se podría brindar

Una vez definidas las intenciones y entidades, se programó el flujo de conversación en Watson Conversation, ver Figura 22, con resultados satisfactorios y operatividad en línea.

[34] IBM Bluemix. Watson conversation API Documentation, About [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://console.bluemix.net/docs/services/conversation/index.html#about.
[35] Ibíd.

Figura 22. Conversaciones en Chatbot de Watson Analytics mencionando problemas de VPN y solicitud de instalación de aplicativo



Fuente: Plataforma de Watson Conversation en Bluemix.

14.4. ANÁLISIS DE LOS RESULTADOS FINALES MEDIANTE LA TEORIA DE COLAS Y SIMULACIÓN EN EL SOFTWARE ARENA

Una vez mitigado el encolamiento de la mesa de ayuda con la implementación de los dos softwares de inteligencia computacional, se debe analizar qué optimización de costos se podría lograr, para ello es necesario recopilar la cantidad de tickets que se podrían haber reducido de la herramienta de gestión (MAXIMO), ver Tabla 17.

Dado que 871 tickets en el mes de enero se hubiesen podido evitar, se debe hallar la diferencia en la cantidad de llamadas entrantes (según del distribuidor automático de llamadas), posteriormente identificar la tasa media de llamadas por hora, ver Tabla 18, utilizar la información para calcular el número adecuado de servidores mediante la teoría de colas y de esta forma optimizar los costos.

Tabla 17. Cantidad de tickets según el tipo de plataforma donde se optimizó la solución del chatbot y el software de diagnóstico de problemas

Plataforma	Tickets
Internet Explorer	295
Outlook	261
Lync	68
VPN UAG	57
Windows 8	86
Equipo Completo	38
Windows	34
Windows 7	32
Total	871

Fuente: Elaboración del autor.

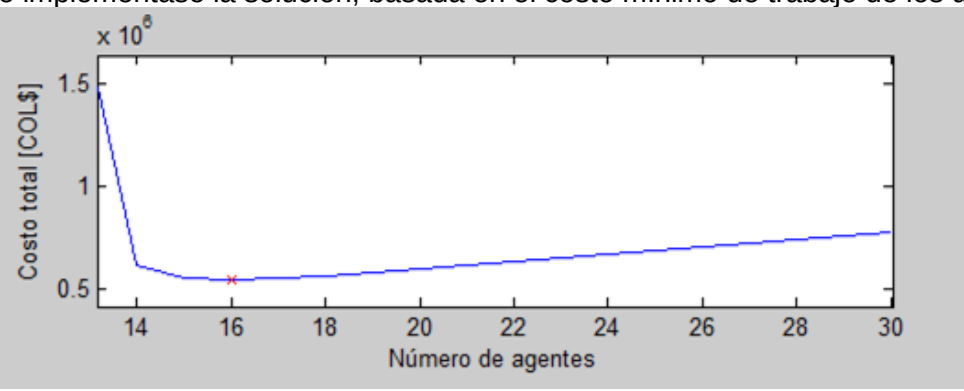
Tabla 18. Comparación de la totalidad de llamadas entrantes, media diaria y media por hora de la situación actual de la mesa de ayuda y su excepción según plan de mejora.

Descripción	TOTAL	MEDIA DIARIA	MEDIA POR HORA
Llamadas entrantes en el mes	5749	302.5789	37.8224
Llamadas entrantes en el mes con excepción	4878	256.7368	32.0921

Fuente: Elaboración del autor.

Dado que la tasa de media de llamas entrantes por hora reduciría, los costos por operación también debido a que se necesitaría menos agentes para atender las necesidades del cliente. Adicionalmente no se debe olvidar que para los costos de espera se debe tener en cuenta que el costo de trabajo del usuario afectado es un factor determinante, por ello se realizará el ejercicio nuevamente con un salario base y con uno de ejecutivo y de esta forma optimizar la Ecuación 13.

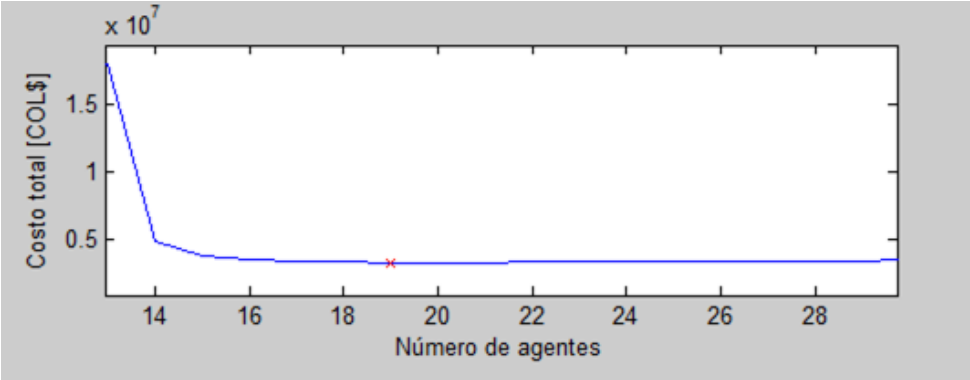
Figura 23. Variación Costo total vs cantidad de servidores, en la situación hipotética donde se implementase la solución, basada en el costo mínimo de trabajo de los usuarios.



Fuente: Elaboración del autor.

Si la totalidad de usuarios que consultan la mesa de ayuda tuviesen un salario básico, se necesitaría contratar a 16 agentes, ver Figura 23, con un costo de servicio y de espera reducido a COP$546,810.00 por hora y de COP$83,115,120.00 por mes, una diferencia y ahorro de COP$13,874,560.00 mensuales con respecto a la situación inicial, ver Figura 15.

Figura 24. Variación Costo total vs cantidad de servidores, en la situación hipotética donde se implementase la solución, basada en el costo máximo de trabajo de los usuarios.



Fuente: Elaboración del autor.

Para el caso en el que los usuarios tuviesen un rango salarial más alto, el costo de espera incrementaría, por ende, se tendría que contratar a 19 agentes, ver Figura 24, con un costo de servicio y de espera reducido a COP$259,890.00 por hora y de COP$39,503,280.00 por mes, una diferencia y ahorro de COP$554,391,120.00 mensuales respecto a la situación inicial, ver Figura 16.

Para obtener los indicadores de calidad de la mesa de ayuda, se utilizará el modelamiento de los usuarios en el software de simulación arena para validar qué comportamiento tendría el sistema si se contrataran 4 agentes de mesa adicionales, para un total de 16 y se redujera la tasa de entrada debido a la implementación del chatbot y el software de diagnóstico.

Figura 25. Cantidad de llamadas atendidas y abandonas con 16 agentes, con software de simulación Arena.

Fuente: Resultados del software de simulación Arena.

Para definir el tiempo de atención de los 4 agentes adicionales, se halló la media de los 12 actuales, ver Tabla 9. Una vez realizada la simulación para los 19 días hábiles de enero y obtener los resultados, ver Figura 25, se utilizó la Ecuación 1 y la Ecuación 5 para obtener la tasa de disponibilidad y de abandono proyectadas, ver Tabla 19.

Tabla 19. Comparación de los indicadores de calidad de la mesa de ayuda con 16 agentes e implementando el chatbox y software de diagnóstico.

Entrantes	Contestadas	Sin Atender	Tasa de Disponibilidad		Tasa de Abandono	
			Simulación	Target	Simulación	Target
4916	4610	259	***94%***	99.96%	***5%***	5.00%

Fuente: Elaboración del autor.

Con respecto a la situación inicial de los indicadores de calidad, ver Tabla 2, aumentaría la tasa de disponibilidad un 16% y la tasa de abandono se reduciría proporcionalmente un 16%. A pesar de que faltaría aumentar la tasa de disponibilidad para cumplir las necesidades del cliente, es necesario tener en cuenta que la mesa de ayuda está pasando por un proceso de transición, es decir que se acabó de implementar, en toda la fase se presenta una variación en la cantidad de consultas a la mesa, donde pasados 6 meses, se llegaría a una estabilización y una reducción abismal en la utilización del servicio.

La rentabilidad del proyecto y la satisfacción del cliente no son los únicos factores que se vieron beneficiados, sino también el hecho de haber implementado dos herramientas de inteligencia computacional con un alto nivel de interacción entre el hombre y la máquina, debido a que en el caso del chatbot se obtenía una conversación fluida con el objetivo final de instruir al usuario. En el caso del sistema de diagnóstico entrenado con redes neuronales, el usuario también podía ser parte de ese entrenamiento y así mismo se creaba un ambiente cooperativo entre las dos partes.

Estas características son parte de la computación cognitiva, que se refiere a sistemas que aprenden a escala, razonan con propósito e interactúan con los humanos de forma natural. En lugar de estar explícitamente programados, aprenden y razonan de sus interacciones con nosotros y de sus experiencias con su entorno.[36]

> Nada de esto implica sensibilidad o autonomía por parte de las máquinas. Más bien, consiste en aumentar la capacidad humana para comprender y actuar sobre los sistemas complejos de nuestra sociedad. Esta inteligencia aumentada es el siguiente paso necesario en nuestra capacidad para aprovechar la tecnología en la búsqueda del conocimiento, para mejorar nuestra experiencia y mejorar la

[36] KELLY III, John E. Computing, cognition and the future of knowing How humans and machines are forging a new age of understanding. Somers NY: IBM Global Services. 2016. p.2.

> condición humana. Es por eso, que representa no solo una nueva tecnología, sino el comienzo de una nueva era de tecnología, negocios y sociedad: la Era Cognitiva.[37]

Es cierto que los sistemas cognitivos son máquinas inspiradas por el cerebro humano. Pero también es cierto que estas máquinas inspirarán al cerebro humano, aumentarán nuestra capacidad de razonar e reformularán las formas en que aprendemos. En el siglo XXI, conocer todas las respuestas no distinguirá la inteligencia de alguien, sino que la capacidad de hacer mejores preguntas será la marca del verdadero genio.[38]

[37] KELLY III, John E. Computing, cognition and the future of knowing. Op. cit., p.2.
[38] Ibíd., p.11.

15. CONCLUSIONES.

-Para realizar una óptima gestión administrativa en una mesa de ayuda en primera instancia se debe validar la capacidad del servicio, que podrá ser analizada mediante el modelamiento de la teoría de colas y que posteriormente podrá ayudar a discernir entre un aumento o disminución de los recursos de la compañía.

-Uno de los resultados que le dan mayor valor a la compañía es la satisfacción de cada uno de sus trabajadores, que se logra evaluando su capacidad media laboral e implementando un plan de solución donde se construya un sistema que logre potenciar las habilidades como persona y acoplarlas a un grupo de trabajo, sin tener que recurrir a la sobrecarga laboral.

-La implementación de aplicaciones software especializadas en la solución de problemas y consultas es una alternativa que permite mitigar el encolamiento en una mesa de ayuda debido a que evita que el usuario ingrese al sistema de espera.

-Los costos de una mesa de ayuda se deben desglosar en dos tipos, costos de operación y costos de espera, los cuales varían exponencial al modificar la capacidad del servicio, por ende, es necesario encontrar el punto de optimización donde se encuentre el balance entre la calidad del servicio y la rentabilidad de la compañía.

- La implementación de aplicaciones desarrolladas con técnicas de inteligencia computacional, que interactúan con el usuario de manera natural, inspira al hombre a desarrollar su capacidad de razonamiento y estimula su curiosidad por descubrir la frontera de su alcance, conservando así su esencia, el conocimiento.

.

16. BILIOGRAFÍA

PONCE CRUZ, Pedro. Inteligencia Artificial con aplicaciones a la ingeniería. 1 ed. México D.F: Alfaomega Grupo Editor, 2010.

HILLIER, Frederick S; LIEBERMAN, Gerald J. Introducción a la investigación de operaciones. 9 ed. Mexico DF: McGraw-Hill/Interamericana editores, 2010.

KRAJEWSKI, Lee. Operations Management. 10 ed. Kendallville: Pearson Education Limited, 2013.

Haren, Van. Operación del Servicio, Basada en ITIL®V3. 1 ed. Van Haren Publishing, Zaltbommel, 2008.

ESPINOSA, Jorge. Un sistema basado en redes neuronales artificiales para diagnóstico de anemia ferropénica. REVISTA DE INVESTIGACIÓN DE SISTEMAS E INFORMÁTICA, 2012. ISSN 1816-3823.

DEMUTH, Howard. Neural Network Toolbox™. 6 ed. Natick: The MathWorks, Inc. 2009.

ISASI VIÑUELA, Pedro. Redes Neuronales Artificiales, un Enfoque Práctico. 1 ed. Madrid: Pearson Educación S.A, 2004.

IBM. *Workshop - ITIL Foundations V3. Mexico DF: IBM Corporation, 2012.*

IBM. IBM has obtained Corporate wide certifications for ISO 9001, ISO 14001, ISO 50001 and OHSAS 18001 [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www-935.ibm.com/services/us/en/it-services/iso-management-system-certifications.html.

ITIL Central. In A Nutshell: A Short History of ITIL [en línea], [revisado de 2 Octubre de 2017]. Disponible en internet: http://itsm.fwtk.org/History.htm

IBM Bluemix. Watson conversation API Documentation, About [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://console.bluemix.net/docs/services/conversation/index.html#about.

AVAYA. Una reseña sobre nuestra compañía, Historia ITIL [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.avaya.com/es/about-avaya/our-company/company_overview/company-overview/

WALCHUK, Zach. Build a chatbot in ten minutes with Watson [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.ibm.com/blogs/watson/2016/12/build-chat-bot/

IBM Developer Cloud. API Reference, Watson Conversation, Introduction [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.ibm.com/watson/developercloud/conversation/api/v1/?cm_mc_uid=30613969409014902027442&cm_mc_sid_50200000=1501388877&cm_mc_sid_52640000=1501388877#introduction

IBM Cloud Platform. IBM Bluemix [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.ibm.com/cl-es/marketplace/cloud-platform

IBM Market Place. IBM Watson Analytics, What can I do for Business [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.ibm.com/ms-en/marketplace/watson-analytics

AVAYA. Avaya Aura® Call Center Elite [en línea], [revisado 2 de Octubre de 2017]. Disponible en internet: https://www.avaya.com/es/producto/avaya-aura-call-center-elite/

MATHWORKS. Patternnet, documentation [en linea], [revisado 2 de Octubre de 2017]. Disponible en Internet: https://es.mathworks.com/help/nnet/ref/patternnet.html

RUDRAKSHI, Charu. API-fication, Core Building Block of the Digital Enterprise [en linea], Agosto de 2014 [revisado 2 de Octubre de 2017]. Disponible en Internet: http://www.hcltech.com/sites/default/files/apis_for_dsi.pdf

KELLY III, John E. Computing, cognition and the future of knowing How humans and machines are forging a new age of understanding. Somers NY: IBM Global Services. 2016.

ITIL Central. In A Nutshell: The Five ITIL Volumes [en línea], [revisado de 2 Octubre de 2017]. Disponible en internet: http://itsm.fwtk.org/Contents.htm

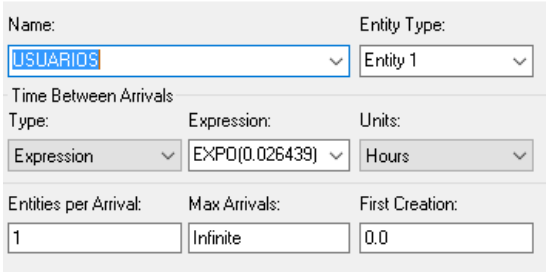
MATHWORKS, Levenberg-Marquardt backpropagation [en línea], [revisado el 2 de Octubre]. Encontrado en internet: https://es.mathworks.com/help/nnet/ref/trainlm.html

OLD DOMINION UNIVERSITY. Evaluating a classification model – What does precision and recall tell me? [en linea], [revisado el 5 de Diciembre]. Disponible en internet: http://www.cs.odu.edu/~mukka/cs795sum10dm/Lecturenotes/Day4/recallprecision.pdf

ANEXOS

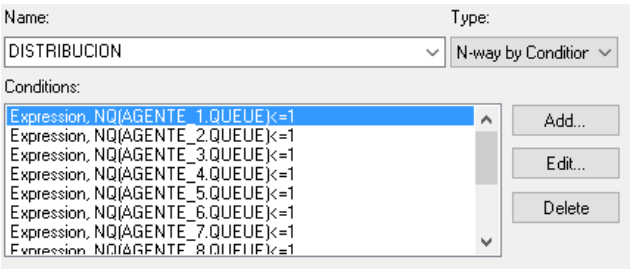
ANEXO A. Configuración y emisión de reportes en la simulación inicial con el software Arena.

Figura 26. Configuración de la tasa de entrada de llamadas en simulación en Arena, con los valores actuales de la mesa de ayuda



Fuente: Simulación Software Arena.

Figura 27. Configuración del sistema de distribución de usuario en software de simulación Arena.



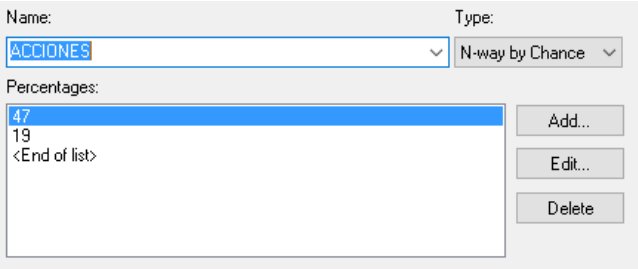
Fuente: Simulación Software Arena.

Figura 28. Configuración de la tasa de atención de cada uno de los usuarios de la mesa de ayuda.

Name	Expression
AGENTE_1	EXPO(0.407506702412869)
AGENTE_2	EXPO(0.375308641975309)
AGENTE_3	EXPO(0.368038740920097)
AGENTE_4	EXPO(0.435530085959885)
AGENTE_5	EXPO(0.434285714285714)
AGENTE_6	EXPO(0.393782383419689)
AGENTE_7	EXPO(0.444444444444444)
AGENTE_8	EXPO(0.361904761904762)
AGENTE_9	EXPO(0.411924119241192)
AGENTE_10	EXPO(0.349425287356322)
AGENTE_11	EXPO(0.438040345821326)
AGENTE_12	EXPO(0.4)

Fuente: Simulación Software Arena.

Figura 29. Configuración de la probabilidad de que una llamada sea solucionada por los diferentes grupos de trabajo en software Arena.



Fuente: Simulación Software Arena.

Figura 30. Reporte de resultados de la simulación de la mesa de ayuda.

	Waiting Time
AGENTE_1.Queue	0.72
AGENTE_10.Queue	0.61
AGENTE_11.Queue	0.72
AGENTE_12.Queue	0.66
AGENTE_2.Queue	0.72
AGENTE_3.Queue	0.77
AGENTE_4.Queue	0.84
AGENTE_5.Queue	0.81
AGENTE_6.Queue	0.67
AGENTE_7.Queue	0.76
AGENTE_8.Queue	0.60
AGENTE_9.Queue	0.78

Fuente: Simulación Software Arena.

Figura 31. Reporte de resultados de cada uno de los agentes de la mesa de ayuda en la simulación del software Arena.

	Inst Util	Num Busy	Num Sched	Num Seized	Sched Util
AUGUSTO	1,00	1,00	1,00	405,00	1,00
CAMILO	1,00	1,00	1,00	405,00	1,00
CRISTIAN	1,00	1,00	1,00	376,00	1,00
EDITH	1,00	1,00	1,00	342,00	1,00
FABIO	1,00	1,00	1,00	346,00	1,00
GIAN	0,99	0,99	1,00	405,00	0,99
HERMES	0,99	0,99	1,00	360,00	0,99
JOSE	0,98	0,98	1,00	427,00	0,98
JULY	0,96	0,96	1,00	324,00	0,96
LUCAS	0,94	0,94	1,00	389,00	0,94
RAUL	0,93	0,93	1,00	322,00	0,93
RICARDO	0,88	0,88	1,00	315,00	0,88

Fuente: Software de simulación Arena.

ANEXO B. Código de programación en Matlab para calcular la optimización de costos en la mesa de ayuda mediante la teoría de colas

```
%Promedio llamadas día 302.5789 // 256.7368
Pld=256.7368;
%Horas de atención
Hda=8;
%tasa de llegada (hora)
l=Pld/Hda;
%tasa de atención (hora)
u=2.5048;
Cs=18200;
%Rango 18200 hasta 230000
Cw=230000;
limi=ceil(l/u);
lim=30;
for s=1:lim
    p=l/(s*u);
    SmPo=0;
    for n=0:s-1
         SmPo=SmPo+(l/u)^n/factorial(n);
    end
    Po=1/(SmPo+(l/u)^s/factorial(s)/(1-l/(s*u)));
    Lq=Po*(l/u)^s*p/(factorial(s)*(1-p)^2);
    L(s)=Lq+l/u;
    E(s)=Cs*s+Cw*L(s);
end
for i=1:limi
     E(i)=E(limi);
end
[costo agentes]=min(E(:))
s=1:lim;
subplot(2,1,1)
plot(s,L,'c')
xlabel('Número de agentes');
ylabel('Valores de L');
title('Variación de L vs s');
subplot(2,1,2)
plot(s,E)
xlabel('Número de agentes');
ylabel('Costo total [COL$]');
title('Variación Costo total vs cantidad de servidores');
hold on;
plot(agentes,costo,'xr','LineWidth',1,'MarkerSize',5);
```

ANEXO C. Configuración de vectores de entrada y salida para el entrenamiento de la red neuronal.

Figura 32. Asignación de vectores de entrada y salida para el entrenamiento de la red neuornal

	SITUACIÓN 1	SITUACIÓN 2	SITUACIÓN 3	SITUACIÓN 4	SITUACIÓN 5	SITUACIÓN 6	SITUACIÓN 7	SITUACIÓN 8	SITUACIÓN 9	SITUACIÓN 10
	VECTOR DE ENTRADAS									
El equipo no tiene conexión a internet.'	1	1	0	0	0	0	0	0	0	0
El equipo presente demasiada lentitud en cualquier tipo de tarea.'	0	1	1	1	1	1	1	1	1	0
'La navegación en internet es lenta en la mayoría de páginas web.'	1	1	1	0	1	0	1	0	0	0
'No se puede ejecutar ninguna aplicación, el equipo no responde.'	0	0	0	1	1	1	1	1	0	0
'No se encuentran algunos archivos que había almacenado en el disco duro.'	0	0	0	0	0	0	1	1	0	0
'El equipo ha mostrado la pantalla azul en repetidas ocasiones con error crítico.'	0	0	0	0	0	1	1	0	0	0
'El equipo se reinicia momentáneamente sin haber programado ninguna actividad de este tipo.'	0	0	0	0	0	1	1	1	1	0
'Aparecen notificaciones de error cuando se intenta guardar un archivo. '	0	0	0	0	0	0	1	1	0	0
'Se siente un agudo y continuo como si dos placas de metal estuviesen rozando dentro del equipo.'	0	0	0	0	0	0	0	1	0	0
'El equipo presenta altas temperaturas de forma recurrente.'	0	0	0	0	0	0	0	0	1	0
'El equipo no presenta ninguna falla de las anteriormente mencionadas.'	0	0	0	0	0	0	0	0	0	1
	VECTOR DE SALIDAS									
Realizar ipconfig/release, ipconfig/renew, ipconfig/flush'	1	0	0	0	0	0	0	0	0	0
'Reinstalar navegador.'	0	1	0	0	0	0	0	0	0	0
'Reiniciar el computador en modo seguro e instalar actualizaciones.'	0	0	1	0	0	0	0	0	0	0
'Desfragmentar el disco.'	0	0	0	1	0	0	0	0	0	0
'Eliminar programas innecesarios, caché de navegadores y temporales.'	0	0	0	0	1	0	0	0	0	0
'Ampliar memoria RAM'	0	0	0	0	0	1	0	0	0	0
'El computador tiene un malware, descargar software antivirus para ejecutar chequeo.'	0	0	0	0	0	0	1	0	0	0
'Cambiar el disco duro.'	0	0	0	0	0	0	0	1	0	0
Repara el sistema de ventilación.'	0	0	0	0	0	0	0	0	1	0
'Solicitar diagnóstico técnico avanzado.'	0	0	0	0	0	0	0	0	0	1

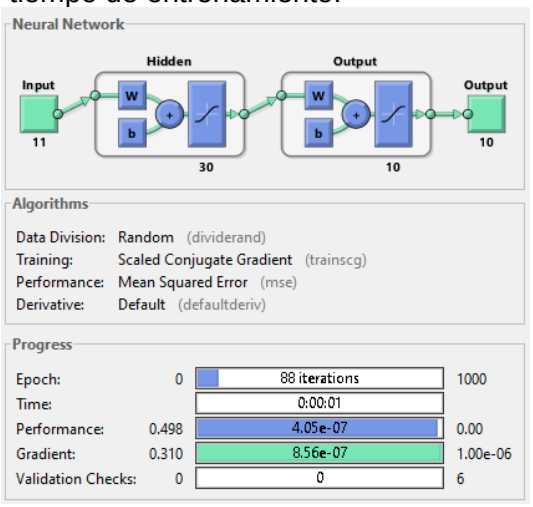
Fuente: Elaboración del autor.

ANEXO D. Código de programación en Matlab para entrenar la red neuronal.

```
load inputs_and_outputs
mired=patternnet(30); %Creación de la red.
mired=train(mired,datos,solucion);%Entrenamiento de la red
save('red.mat','mired');
```

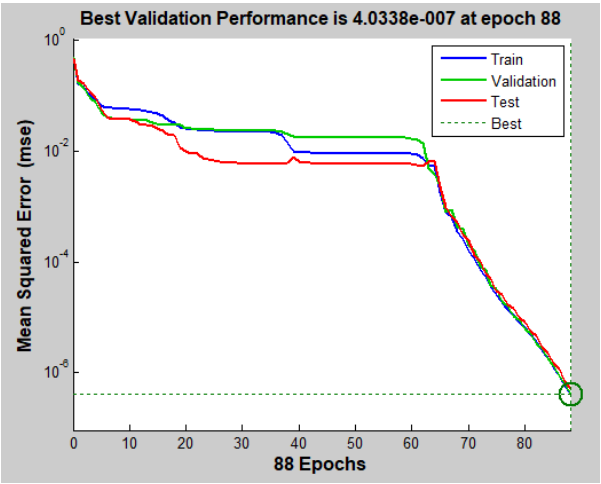
ANEXO E. Resultados del entrenamiento de la red neuronal en Matlab.

Figura 33. Configuración de la red neuronal en Matlab: número de neuronas ocultas, cantidad de iteraciones y tiempo de entrenamiento.



Fuente: Entrenamiento de red neuronal en Matlab.

Figura 34. Cantidad de iteraciones vs media cuadrada del error; basada en el entrenamiento de la red neuronal.



Fuente: Entrenamiento de red neuronal en Matlab.

ANEXO F. Código de programación para la interfaz GUI en Matlab

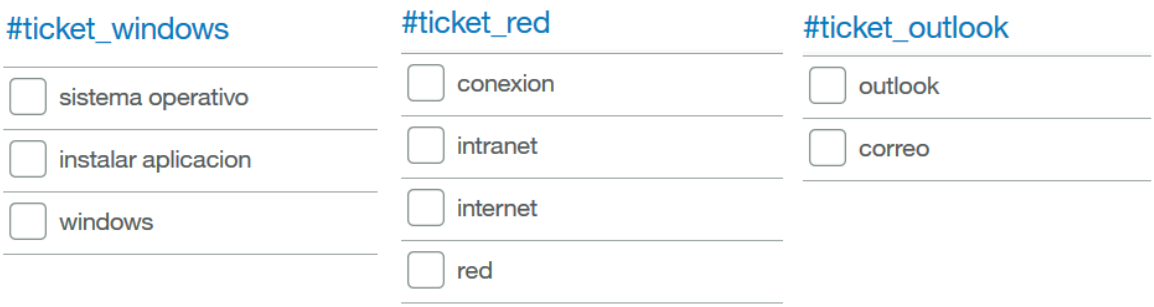
```
Proyecto1ventana()
function pushbutton1_Callback(hObject, eventdata, handles)
clc
A=[0 0 0 0 0 0 0 0 0 0 0];
if(get ( handles.radiobutton1, 'Value'))
    A(1)=1;end
if(get ( handles.radiobutton2, 'Value'))
    A(2)=1;end
if(get ( handles.radiobutton3, 'Value'))
    A(3)=1;end
if(get ( handles.radiobutton4, 'Value'))
    A(4)=1;end
if(get ( handles.radiobutton5, 'Value'))
    A(5)=1;end
if(get ( handles.radiobutton6, 'Value'))
    A(6)=1;end
if(get ( handles.radiobutton7, 'Value'))
    A(7)=1;end
if(get ( handles.radiobutton8, 'Value'))
    A(8)=1;end
if(get ( handles.radiobutton9, 'Value'))
    A(9)=1;end
if(get ( handles.radiobutton10, 'Value'))
    A(10)=1;end
if(get ( handles.radiobutton11, 'Value'))
    A(11)=1;end
select_user=A'
load red
Rec=sim(mired,select_user);
[valor,opcion]=max(Rec);
resp=opcion
save('variables1.mat','resp','select_user')

Proyecto2ventana()
function pushbutton1_Callback(hObject, eventdata, handles)
load variables1
load red
load inputs_and_outputs
selection=get(get(handles.uipanel1,'SelectedObject'),'string');
switch selection
    case 'SI'
            aux=select_user;
            aux2=[0 0 0 0 0 0 0 0 0 0];
            aux2(resp)=1;
            [filas,columnas]=size(datos);
            for j=1:11
             for i=1:11
                 datos(i,columnas+j)=aux(i);
             end
             for i=1:10
                 solucion(i,columnas+j)=aux2(i);
             end
            end
```

```
            mired=patternnet(30); %Creación de la red.
            mired=train(mired,datos,solucion);%Entrenamiento de la red
            save('red.mat','mired');
            save('inputs_and_outputs.mat','datos','solucion')
            h   =   msgbox('¡Gracias   por   su   ayuda,   la   red   se   ha
retroalimentado!', 'NOTIFICACIÓN','Warm');
     case 'NO'
         A=[0 0 0 0 0 0 0 0 0 0 0];
        if(get ( handles.radiobutton3, 'Value'))
            A(1)=1;end
        if(get ( handles.radiobutton4, 'Value'))
            A(2)=1;end
        if(get ( handles.radiobutton5, 'Value'))
            A(3)=1;end
        if(get ( handles.radiobutton6, 'Value'))
            A(4)=1;end
        if(get ( handles.radiobutton7, 'Value'))
            A(5)=1;end
        if(get ( handles.radiobutton8, 'Value'))
            A(6)=1;end
        if(get ( handles.radiobutton9, 'Value'))
            A(7)=1;end
        if(get ( handles.radiobutton10, 'Value'))
            A(8)=1;end
        if(get ( handles.radiobutton11, 'Value'))
            A(9)=1;end
        if(get ( handles.radiobutton12, 'Value'))
            A(10)=1;end
            aux=select_user
            aux2=A'
             [filas,columnas]=size(datos);
            for j=1:11
             for i=1:11
                 datos(i,columnas+j)=aux(i);
             end
             for i=1:10
                 solucion(i,columnas+j)=aux2(i);
             end
            end
            mired=patternnet(30); %Creación de la red.
            mired=train(mired,datos,solucion);%Entrenamiento de la red
            save('red.mat','mired');
            save('inputs_and_outputs.mat','datos','solucion')
            h   =   msgbox('¡Gracias   por   su   ayuda,   la   red   se   ha
retroalimentado!', 'NOTIFICACIÓN','Warm');
end
```

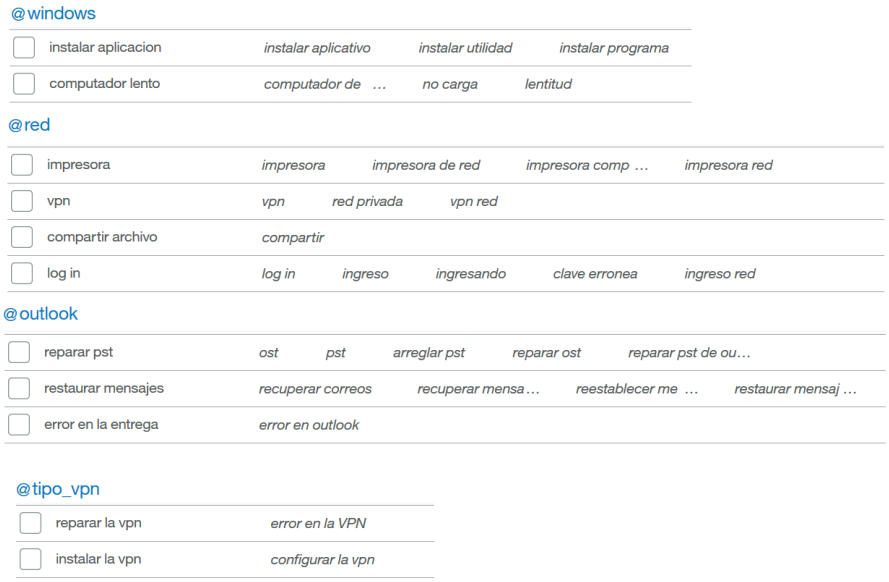
ANEXO G. Diseño, implementación y resultados del chatbot de Watson Conversation en la mesa de ayuda.

Figura 35. Configuración de intentos en el flujo de conversación.



Fuente: Elaboración del autor.

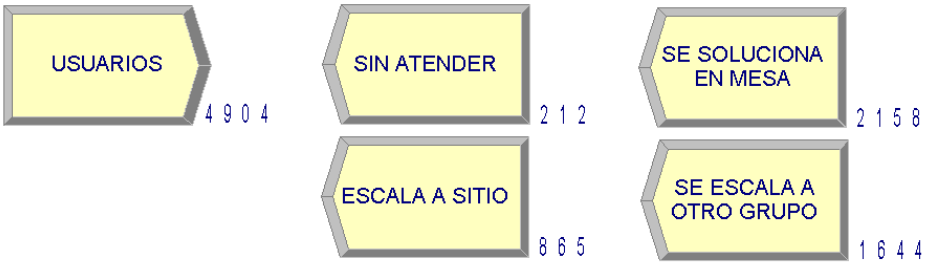
Figura 36. Configuración de entidades en el flujo de la conversación.



Fuente: Elaboración del autor.

ANEXO H. Resultados finales implementando el Chatbox de Watson y el software de diagnóstico de problemas

Figura 37. Resultados del modelamiento de la mesa de ayuda de IBM si se tuviesen 17 agentes.



Fuente: Resultados obtenidos en simulación con software Arena.

Tabla 20. Comparación de los indicadores de calidad de la mesa de ayuda en caso de implementar el chatbot de Watson y el software de diagnóstico, basado en la simulación Arena.

Entrantes	Contestadas	Sin Atender	Tasa de Disponibilidad		Tasa de Abandono	
			Simulación	Target	Simulación	Target
4904	4667	212	***95%***	99.96%	***4%***	5.00%

Fuente: Elaboración del autor.

Figura 38. Reporte de resultados de los tiempos de espera con la ayuda de 17 agentes en la mesa de ayuda.

Time Units: Hours	Waiting Time
AGENTE_1.Queue	0.76
AGENTE_10.Queue	0.50
AGENTE_11.Queue	0.65
AGENTE_12.Queue	0.48
AGENTE_13.Queue	1.44
AGENTE_14.Queue	2.66
AGENTE_15.Queue	3.31
AGENTE_16.Queue	3.45
AGENTE_17.Queue	4.20
AGENTE_2.Queue	0.74
AGENTE_3.Queue	0.69
AGENTE_4.Queue	0.80
AGENTE_5.Queue	0.73
AGENTE_6.Queue	0.76
AGENTE_7.Queue	0.71
AGENTE_8.Queue	0.59
AGENTE_9.Queue	0.63

Fuente: Software de simulación Arena.