
Explicación Tabla de Requerimientos

Requerimientos del Negocio:

- **R1. Obtener una comprensión más completa de los comportamientos de los clientes no regulados o libres.**
 - **Prueba Prevista.** Análisis descriptivo detallado de los clientes.
 - **Criterio o métrica de evaluación y rangos deseados.** 100% de los clientes deben estar cargados con su respectivo resumen descriptivo del comportamiento histórico.
 - **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** 100% de los clientes están cargados, hay un resumen descriptivo del comportamiento histórico en el primer Dashboard y en la serie de tiempo es posible visualizar las cuatro variables estudiadas.
 - **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** La solución tiene cargados los 30 clientes, con toda la data asociada. En el Dashboard 1 hay también una tabla con el resumen descriptivo de consumos y el Dashboard 2 permite visualizar las cuatro variables estudiadas: Energía Activa, Energía Reactiva, Voltaje_FA y Voltaje_FC como una serie de tiempo.
 - **¿Qué ajustes fueron implementados para lograrlo?** Carga de toda la data, creación de visualizaciones que permiten observar el comportamiento histórico de cada variable de estudio y además, filtros que se pueden aplicar para segmentar la información.
 - **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** Si el cliente considera que es importante visualizar más detalles, sin tener en cuenta clusters, es posible crear una tabla adicional con valores específicos de las cuatro variables de estudio relacionados con cuartiles para analizar de otra manera el comportamiento de clientes y sectores. Asimismo, es posible agregar diagramas de cajas y bigotes o frecuencias.
 - **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Agregar una hoja adicional con detalles específicos del análisis descriptivo como media, mediana, valor mínimo y valor máximo para cada cliente, y visualizarlo con información de lecturas con anomalías para facilitar el trabajo de los analistas de pérdidas.
- **R2. Identificar anomalías asociadas a pérdidas no técnicas.**
 - **Prueba Prevista.** Análisis de clústers.
 - **Criterio o métrica de evaluación y rangos deseados.** Reconocer clústers con comportamientos anómalos asociados a pérdidas no técnicas.
 - **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Dashboard con clústers con comportamientos anómalos asociados a pérdidas no técnicas, teniendo en cuenta la regla creada que relaciona la Energía Activa y Reactiva.
 - **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** Definición de regla para anomalías considerando lecturas no típicas, específicamente casos donde la energía reactiva es significativamente mayor que la energía activa pues esperamos que la energía activa sea mayor o, al menos, no mucho menor que la energía reactiva. Después, desarrollo de un modelo de clustering para la segmentación de los clientes, elegido entre diferentes propuestas y generación de visualizaciones que permiten reconocer diferentes clusters con sus características específicas.
 - **¿Qué ajustes fueron implementados para lograrlo?** Creación de regla para reconocer posibles anomalías teniendo en cuenta una relación entre la energía activa y reactiva.
 - **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** Esta regla es definida teniendo en cuenta el análisis descriptivo de la data y la bibliografía asociada al caso de estudio, por lo tanto, el analista de pérdidas debe analizar el prototipo para obtener insights.
 - **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** El modelo de clustering y la regla propuesta, no define con exactitud una anomalía pero si una posible lectura anómala que debe ser estudiada, es importante que el equipo de analítica reciba data adicional con una variable específica de anomalías que permita desarrollar un

modelo más robusto.

- **R3. Comparar información de los clientes por sector que permita identificar tendencias.**
 - **Prueba Prevista.** Análisis descriptivo de los clientes por sector.
 - **Criterio o métrica de evaluación y rangos deseados.** Presentar resumen descriptivo por sector.
 - **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Resumen descriptivo del comportamiento histórico en el primer Dashboard y en la serie de tiempo visualización de las cuatro variables estudiadas, con segmentación por sector.
 - **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** En el Dashboard 1 hay un análisis descriptivo con información de clústers definidos teniendo en cuenta la regla que relaciona Energía Activa y Reactiva planteada anteriormente, asimismo presentación de serie de tiempo de las cuatro variables de estudio. Todo esto, con la posibilidad de la segmentación por sector por medio de filtros que permite un análisis diferencial para identificar tendencias.
 - **¿Qué ajustes fueron implementados para lograrlo?** Principalmente, aplicación de filtros que permiten la segmentación de la data.
 - **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** Es posible segmentar todas las visualizaciones por sectores y además por clientes dentro de ese sector, sin embargo, si el usuario desea un análisis diferente con anomalías reconocidas, es necesario obtener la data adicional.
 - **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Agregar label de anomalías con data adicional para hacer un análisis específico de anomalías reales por sector para definir nuevas reglas en el modelo.

Requerimientos de Desempeño:

- **R4. Análisis de clústers para identificar tendencias en las agrupaciones realizadas por el algoritmo y determinar qué factores influyen en la presencia de comportamientos anómalos.**
 - **Prueba Prevista.** Métrica de desempeño análisis de clústers.
 - **Criterio o métrica de evaluación y rangos deseados.** Coeficiente de Silhouette > 0.4 .
 - Esto sugiere que los clústers están relativamente bien definidos y separados entre sí.
 - **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Coeficiente de Silhouette > 0.4 . Esto sugiere que los clústers están relativamente bien definidos y separados entre sí.
 - **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** Desarrollo de diferentes modelos de clustering como DBSCAN y KMeans para segmentar los datos y detectar posibles anomalías. Para empezar con este modelo, definimos una regla de aplicación que considera la información relevante que encontramos durante el análisis de la base de datos y que posiblemente representa una anomalía relacionando la Energía Activa y Reactiva. El análisis de resultados para seleccionar el mejor algoritmo de clustering fue realizado comparando los diferentes valores del Coeficiente de Silhouette, en este caso, KMeans resulta en un coeficiente igual a 0.6.
 - **¿Qué ajustes fueron implementados para lograrlo?** El modelo fue calibrado y el resultado sugiere la segmentación de la data en 4 clusters.
 - **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** En caso de que el cliente sugiera una regla o agrupación diferente, el equipo deberá desarrollar un nuevo modelo.
 - **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Incluir los labels de las diferentes lecturas para crear un modelo que permita agrupar teniendo en cuenta esas anomalías ya probadas.
- **R5. Seleccionar mejor algoritmo que permita identificar anomalías.**
 - **Prueba Prevista.** Métricas de desempeño.
 - **Criterio o métrica de evaluación y rangos deseados.** Coeficiente de Silhouette más

cercano a 1 para algoritmos como PCA, LOF y DBSCAN.

- **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Selección de modelo de aprendizaje no supervisado Isolation Forest, comparando Coeficiente de Silhouette con otros modelos como Local Outlier Factor.
- **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** Desarrollo y calibración de diferentes modelos para detectar anomalías, incluyendo PCA, Local outlier Factor, DBSCAN, One Class SVM y también Isolation Forest. Después, análisis de resultados comparando Índice Davies-Bouldin y Coeficiente de Silhouette.
- **¿Qué ajustes fueron implementados para lograrlo?** Investigación de diferentes modelos para detectar anomalías en una serie de tiempo considerando outliers, además, desarrollo, calibración, implementación y evaluación de los diferentes algoritmos para seleccionar el mejor para la data de estudio.
- **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** Si las posibles anomalías detectadas nos son positivos verdaderos, es necesario recalibrar el modelo con los labels reales.
- **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Incluir labels reales que debe proporcionar ElectroDunas, para calibrar el modelo.

- **R6. Desarrollar un modelo de aprendizaje no supervisado que permita identificar anomalías.**

- **Prueba Prevista.** Métricas basadas en la densidad.
- **Criterio o métrica de evaluación y rangos deseados.** Métrica de Densidad Local y Densidad Local Promedio.
- **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Coeficiente de Silhouette más cercano a 1, desarrollo de algoritmo calibrado Isolation Forest para detección de anomalías.
- **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** Las métricas arrojadas por los algoritmos son consideradas aceptables para el clustering, por KMeans el coeficiente de silueta resultó en 0.6 y para el algoritmo de Isolation Forest 0.37, lo que indica una separación equitativa entre los grupos.
- **¿Qué ajustes fueron implementados para lograrlo?** Fue necesario calibrar hiperparámetros en ambos algoritmos, como el número de clusters, el método de inicialización y número de estimadores.
- **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** La presencia de outliers o falta de data que diferencie los grupos son posibles causas de un mal agrupamiento, por lo que esto deberá ser tratado en la etapa de procesamiento de datos y socializado con el cliente, ya que mediciones incorrectas afectarán el desempeño de los algoritmos.
- **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Recolectar más datos de los clientes hará que los algoritmos tengan más información y por ende brinden mejores resultados. También recomendamos aclaración en cuanto a lo que es considerado una anomalía, ya que los criterios pueden ser distintos entre ambas partes.

Requerimientos Funcionales:

- **R7. Construir un dashboard amigable con el usuario, con una interfaz clara y comprensible.**

- **Prueba Prevista.** Demo a usuarios.
- **Criterio o métrica de evaluación y rangos deseados.** Dashboard que cumple con requerimientos.
- **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Dashboard amigable y presentación en demo para los usuarios.
- **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** La selección de gráficos e indicadores fue pensada para un público no experto en el tema, el cual fácilmente podrá interpretar los resultados del modelo y utilizar el dashboard de una manera simple e intuitiva.
- **¿Qué ajustes fueron implementados para lograrlo?** Fue planteado un modelo relacional que permite mostrar la data de manera sencilla y seleccionamos gráficas de uso común, como gráficas de dispersión y líneas temporales de fácil interpretación.

-
- **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** El equipo deberá replantear el diseño del dashboard en caso de que el cliente no lo considere amigable y práctico para operación.
 - **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Sugerimos al cliente brindar constante feedback de la herramienta para poder generar mejoras acorde a sus necesidades.
 - **R8. Requiere licencia de Power BI.**
 - **Prueba Prevista.** Revisión.
 - **Criterio o métrica de evaluación y rangos deseados.** Cumple.
 - **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Licencia Power BI, cumple.
 - **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** La licencia de Power BI es un requerimiento que el cliente debe adquirir, su precio oscila entre 10USD y 20 USD mensuales por usuario.
 - **¿Qué ajustes fueron implementados para lograrlo?** En este caso, la Universidad proporciona esta licencia, por lo que no fue requerido ningún proceso externo para desarrollar el proyecto.
 - **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** En caso de que Power BI no cumpla con los requerimientos del cliente, es indispensable evaluar otra herramienta de visualización, como Looker Studio o Tableau
 - **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Contabilizar el número de usuarios que usaran la aplicación para no incurrir en sobrecostos y definir si Power BI cumple con los requerimientos operacionales.
 - **R9. Cargar el dashboard rápidamente.**
 - **Prueba Prevista.** Pruebas de Performance.
 - **Criterio o métrica de evaluación y rangos deseados.** Cargar front end en menos de 10 segundos.
 - **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Cargar front end en menos de 10 segundos.
 - **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** El dashboard fue disponibilizado en el siguiente [link](#), donde su tiempo de carga una vez logueado es inferior a 10 segundos.
 - **¿Qué ajustes fueron implementados para lograrlo?** La implementación de gráficas funcionales y el correcto planteamiento del modelo relacional hace que los tiempos de carga y actualización sean reducidos, por lo que el correcto planteamiento de esto es la clave.
 - **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** Es posible que el requerimiento no sea cumplido debido a factores externos, como conexión a internet o velocidad de los dispositivos donde es ejecutado el dashboard, por lo que esto deberá aclararse en el momento de su disponibilización.
 - **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** En caso de contemplarse la inclusión de más datos o elementos en el tablero, el cliente deberá considerar la afectación de tiempos de carga y funcionamiento.
 - **R10. Aplicar filtros de fecha, hora, cliente y sector económico para visualizar subgrupos específicos.**
 - **Prueba Prevista.** Revisión.
 - **Criterio o métrica de evaluación y rangos deseados.** Cumple.
 - **Criterio o métrica de evaluación y rangos deseados, SOLUCIÓN ENTREGADA.** Filtros disponibles para segmentar por clientes, fecha y sectores principalmente, cumple.
 - **¿Cómo se satisfacen cada uno de los requerimientos y métricas de evaluación definidos?** Fueron integraron slicers en el tablero que permiten la segmentación según los requerimientos definidos.
 - **¿Qué ajustes fueron implementados para lograrlo?** La herramienta de Power BI cuenta

con la funcionalidad de agregar filtros en el tablero, por lo que únicamente fue necesario plantear un modelo relacional de datos de la manera correcta para obtener un filtrado preciso.

- **¿Cómo justificar los casos en que no se satisfaga algún requerimiento?** En caso de que la segmentación no cumpla con los requerimientos, el equipo deberá indagar si las gráficas utilizadas son las correctas, o si el requerimiento del cliente es realizable en el tablero, ya que las herramientas de visualización tienen sus limitaciones.
- **¿Qué acciones correctivas se proponen al cliente para esta o próximas iteraciones?** Recolectar más campos de información de los clientes permitirá realizar un filtrado más preciso y llegar a conclusiones más asertivas.