

# El Big Data aplicado en la industria 4.0: un caso en el sector textil colombiano con un enfoque en la inteligencia de negocios

*Big Data applied in industry 4.0: a case in the Colombian textile sector with a focus on business intelligence*

Gustavo Andrés Araque González<sup>1</sup>, Víctor José Giampietro Torres<sup>2</sup>

**Tipo de Artículo:** Investigación revisión.  
**Recibido:** 13/07/2022 **Aprobado:** 6/09/2 **Publicado:** 18/12/2022

**Resumen:** La competitividad industrial y la evolución operacional han generado la necesidad de desarrollar nuevos e innovadores métodos de sistematización industrial. Como resultado de lo anterior, la presente investigación, de carácter mixto, propone el desarrollo de un modelo de producción industrial enfocado en la industria 4.0 (Big Data) para pymes de confección, con el fin de mejorar los procedimientos, puestos de trabajo y costos, a través de las siguientes fases: 1) *recopilación de las informaciones* de las bases de datos; 2) *limpieza de bases de datos* y correcta edición de informaciones; 3) *modelado de datos* e interrelación de las variables en las bases de datos; 4) *visualización gráfica* de datos (*dashboard*), apoyado en software Power BI, en la visualización

y análisis de los datos, y 5) *análisis y toma de decisiones*. Los resultados obtenidos permitieron generar una mejora del 20% en la identificación de fallas operacionales y no operacionales del proceso industrial interno.

**Palabras clave:** Big Data, ciencia de datos, análisis de datos, sector textil.

**Abstract:** Industrial competitiveness and operational evolution have generated the need to develop new and innovative methods of industrial systematization. As a result of the above, this mixed research proposes the development of an industrial production model focused on industry 4.0 (Big Data) for clothing SMEs, for the improvement of

<sup>1</sup> Gustavo Andrés Araque González. Institución Universitaria Politécnico Gran Colombiano. Colombia, Correo electrónico: garaque@poligran.edu.co ORCID: 0000-0001-8627-8924.

<sup>2</sup> Víctor José Giampietro Torres. Institución Universitaria Politécnico Gran Colombiano. Colombia, Correo electrónico: vigiampietro@poligran.edu.co

procedures, jobs and costs, through the following phases: 1) collection of information from the databases; 2) cleaning of databases and correct editing of information; 3) modeling of data and interrelation of variables in the databases; 4) graphic data visualization (Dashboard), supported by Power BI software, in data visualization and analysis, and 5) analysis and decision making. The results obtained allowed to generate a 20% improvement in the identification of operational and non-operational failures of the internal industrial process.

**Keywords:** Big Data, data science, data analytics, textile sector

## I. Introducción

En la actualidad, el sector textil es particularmente competitivo en la cadena de producción en línea; sin embargo, uno de los factores clave para la mejora y evolución industrial es el tratamiento tecnológico de las informaciones, conocido también como la ciencia de los datos, el cual, de acuerdo con [1] y [2], se define como la generación de cantidades masivas de datos generados que pueden incluir registros operacionales, registros de fallas operativas y no operativas, registros de compras de los clientes, servicio posventa, control de calidad textil y sus variables, entre otros, recopiladas mediante el uso de dispositivos tecnológicos de monitoreo y control. Además, la creciente nube de información que se genera en Internet, a través de la comunicación entre empresas y público interesado, ha generado la necesidad de darle un mejor manejo al conjunto de datos masivos, así como su tratamiento y sistematización [3].

La industria de pequeñas y medianas empresas (pymes) de confecciones ha generado el estudio y desarrollo de nuevos modelos que pueden utilizar arquitecturas de datos para procesarlos más rápida y eficientemente [4]. Sin embargo, los métodos de análisis disponibles son insuficientes para utilizar datos de alta velocidad que fluyen desde varias fuentes, debido a sus complejidades de bajo nivel y deficiencias. En relación con lo anterior, el

uso de Big Data en la industria 4.0 se enfoca en la aplicación de estudios en ciencia de los datos organizacionales de gran volumen, con el propósito de desarrollar nuevas informaciones de valor agregado para las empresas. Este tipo de datos puede tomar diferentes formas (estructurados, semiestructurados o no estructurados) para el estudio de interés [5], [6].

En la presente investigación se realiza la generación, propuesta y diseño de un modelo de producción industrial con enfoque en Big Data en la toma de decisiones organizacionales para la línea de manufactura de confecciones en pymes. El estudio y aplicación de diferentes metodologías de análisis en la ciencia de los datos se relacionan a continuación:

1) *Recopilación de las informaciones de las bases de datos* con la extracción de las diversas fuentes de información en la nube (Oncloud), bases de datos tipo metadatos .csv y archivos de datos estructurados en Excel, tipo .xls.

2) *Limpieza de bases de datos y correcta edición de informaciones* a través del uso de la herramienta de edición Power Query [7] para la edición y limpieza de las bases de datos analizadas.

3) *Modelado de datos e interrelación de las variables en las bases de datos* a través del Power Pivot en la generación de modelo de datos y generación de la relación de bases de datos.

4) *Visualización de gráfica de datos (dashboard)*, apoyada en software Power BI, en la visualización y análisis de los datos.

5) *Análisis y toma de decisiones.*

Todo lo anterior se realiza a través del caso de estudio en búsqueda de la elaboración de un modelo en tiempo real que permita minimizar el espacio de tiempo y los costos relacionados para la toma de decisiones organizacionales, en pro de aumentar los niveles de productividad y rentabilidad de la empresa.

## II. Materiales y métodos

### Materiales

La literatura existente en [8], [9] y [10] provee una gran variedad de definiciones de pruebas de software.

### Métodos

La presente investigación, en relación con sus fines, [11], [12], [13], puede caracterizarse como de tipo metodológica y aplicada. Se define como metodológica porque sigue un estudio ordenado que considera como base de conocimiento la ciencia de los datos, desde la recolección de datos (bases de datos) y la identificación de escenarios de visualización de datos (Big Data), hasta su estructuración, analítica y generación de indicadores clave de desempeño (Key Performance Indicators - KPI) estandarizados en la toma de decisiones organizacionales de la empresa de estudio. La investigación se define como aplicada debido a la búsqueda de una resolución a un problema específico, en el caso de estudio expuesto, y es la correcta automatización y generación de informaciones en tiempo real de salida (monitor de comando e informaciones KPI en tiempo real). En el diagnóstico presente, la organización cuenta con sistemas informáticos para la obtención y recolección de informaciones; sin embargo, este tipo de informaciones de salida no se presentan como "informaciones inteligentes" que permitan generar una interpretación eficiente y eficaz del proceso industrial investigado. Es de acuerdo con esta necesidad que se plantea el modelo de producción industrial con enfoque en Big Data, que les permita utilizar de forma correcta los datos de salida y así mejorar los procesos y procedimientos y aumentar los niveles de calidad, rentabilidad y servicio con los clientes [14].

En cuanto a los medios de investigación, la propuesta es definida como investigación experimental y estudio de caso. Investigación experimental cuyo propósito es detallar y profundizar el modelo productivo óptimo con enfoque en industria 4.0 (Big Data) para un tamaño de muestra definido para empresas (pymes) del sector de confecciones bajo tres (3)

experimentos principales: análisis de calidad de los procesos industriales, análisis de reprocesos de maquinaria y niveles de producción operacionales. Su característica experimental se argumenta en la manipulación de variables independientes asociadas al proceso productivo y la evolución y variaciones que producen tales manipulaciones y control en las variables dependientes [15], [16].

Con respecto a la naturaleza de la investigación, el presente estudio se clasifica como mixto, es decir, de carácter cualitativo y cuantitativo, ya que el estudio implica el uso de información, como, por ejemplo, bases de datos cuantitativas con informaciones de niveles de ventas, producción, número de defectos, etc. en el modelamiento del diseño en un software especializado [17], [18], y cualitativo porque es analizado a través de pruebas de juicio de expertos (método Delphi) en cuanto a las recomendaciones industriales de los resultados generados. A continuación se describen las etapas de la metodología aplicada.

### 1. Recolección de bases de datos

La metodología de la presente investigación inicia con la recolección de las informaciones estructuradas expuestas en los sistemas informáticos locales de la organización, producto del análisis de variables de estudio, como, por ejemplo, las bases de datos de los reprocesos industriales, producción total, tiempos perdidos no operativos, tandas por máquinas, consumos, entre otras. El modelo propuesto parte de, al menos, 10 bases de datos y 148 variables diferentes.

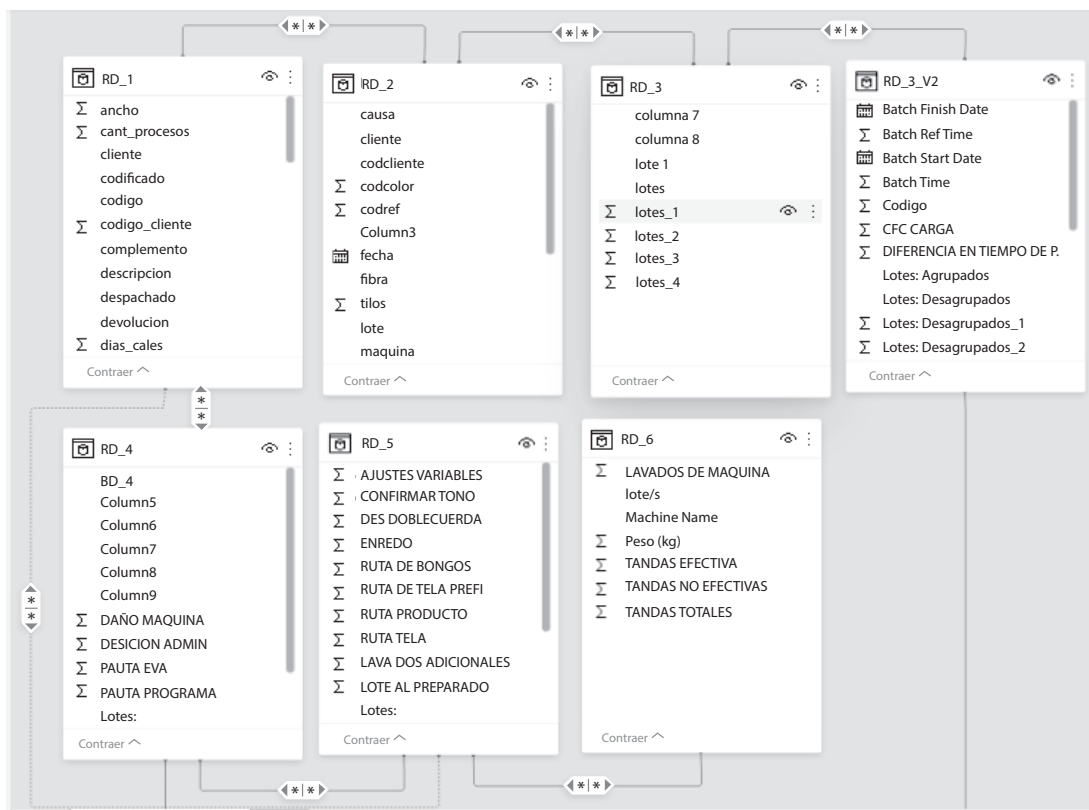
### 2. Definición del modelo relacional y claves principales

El modelo relacional parte del grupo de bases de datos maestras y transaccionales que reflejan la operación de la empresa de estudio, para representar los datos y las relaciones entre ellas. La estructuración de bases de datos parte de una relación "uno a varios", a partir de una clave principal, la cual vinculará los conjuntos de bases de datos a analizar. Para el desarrollo

de esta etapa, fue utilizado el software Power BI Desktop Versión: 2.95.983.0 de Microsoft [19]. En la Figura 1 se observa el modelo relacional

de las bases de datos contenidas en el estudio, donde la clave principal es el número de lote de producción.

**Figura 1.** Modelo relacional de estudio.



Fuente: Elaboración propia (2021)

### 3. Depuración y estandarización de bases de datos

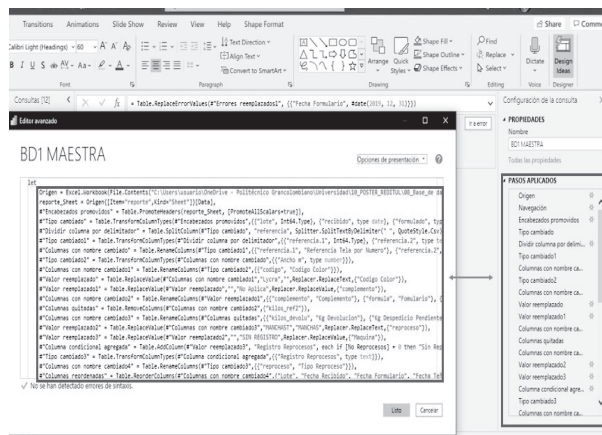
Para elaborar de forma correcta la sincronización y estructuración de la conexión de las bases de datos en tiempo real al interior del software Power BI, es esencial la depuración o limpieza de los datos. Para esta etapa fue utilizada la aplicación Power Query, herramienta que tiene como principal característica permitir importar las informaciones de distintas fuentes de datos, transformar la forma de los datos según sea necesario y luego cargarlos en algún sitio para su uso posterior, ya sea en una

tabla o en un modelo de datos. La herramienta genera de forma automática código en lenguaje M que puede ser ejecutado cada vez que se actualiza una base de datos, garantizando que los datos ingresen al modelo de forma estandarizada.

La Figura 2 ilustra un ejemplo de la edición de la base de datos 1 (BD1 -maestra) del modelo de datos de estudio. En la representación puede observarse, en la parte izquierda, el código de programación generado en Power Query (Lenguaje M): cada vez que se realizan la edición o configuraciones necesarias para el tratamiento del conjunto de

datos, se genera una línea diferente de código, en la parte derecha, en forma de lista. Estas rutas de edición de datos juegan un papel fundamental en la colocación de las informaciones en tiempo real y cuando se trabaja en multiescenarios para la representación de diferentes conjuntos de variables, en su configuración y análisis.

Figura 2. Editor de datos Power Query.



Fuente: Elaboración propia (2021)

#### 4. Analítica de datos

La analítica de datos corresponde al proceso de análisis computacional sistemático de los datos del modelo relacional. El objetivo de esta etapa es descubrir, visualizar, interpretar y comunicar patrones significativos o tendencias generales en los datos. Pueden desarrollarse diferentes tipos de análisis de datos: descriptivo (permite resumir grandes volúmenes de datos en medidas de tendencia central o de variabilidad), predictivo (analiza los datos actuales e históricos reales para hacer predicciones acerca del futuro o acontecimientos no conocidos), prescriptivo (permite encontrar una solución entre una gama de variantes con el objetivo de optimizar los recursos y aumentar la eficiencia operativa; permite planear la asignación de recursos para lograr un objetivo específico) y cognitivo (permite, de la forma más natural posible, tal como aprenden los humanos,

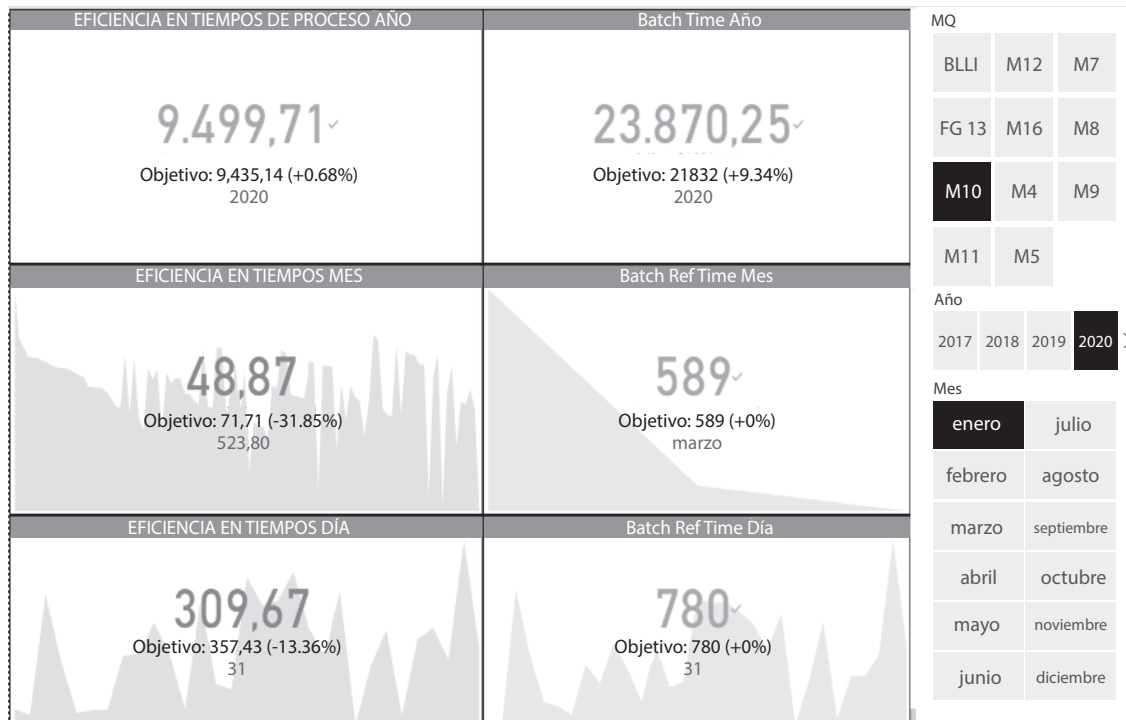
que las tecnologías tengan la habilidad de aprender, razonar y determinar escenarios que respondan a una necesidad). Esta etapa se ha centrado en un análisis de datos descriptivo dentro de las visualizaciones propuestas.

#### 5. Diseño de indicadores clave

Para el diseño de tableros estratégicos (*dashboard*) se diseñaron indicadores KPI clave para la toma de decisiones en el proceso productivo. Fueron utilizadas las bases de datos para la elaboración de cada elemento de indicador: eficiencia de tiempos de procesamiento en año, mes y día; *batch time* o tiempo de lote de producción en año, mes, día; adicionalmente, la meta-objetivo de cada indicador es relacionado de cada resultado de indicador, de forma que se pueda monitorear en tiempo real. Las relaciones y uso de recursos, como materiales, mano de obra y maquinaria por cada lote de producción, determinan la productividad del proceso de teñido de tela y permiten visualizar las fallas operacionales. Otro aspecto fundamental es entender el comportamiento de cada cliente para diseñar modelos de negocio ajustados a la demanda y garantizar que los costos reflejen la operación de teñido de tela y se distribuyan proporcionalmente al consumo de recursos clave.

De acuerdo con las informaciones anteriores y considerando la Figura 3, los indicadores clave de desempeño (KPI) permiten, al interior del caso de estudio, el seguimiento, control y estandarización de las actividades ejecutables en cada una de las fases operativas del modelo industrial de estudio. La Figura 3 expone un ejemplo de indicadores clave de desempeño, aplicado a la maquina industrial 10 (M10) de proceso de teñido a color, en relación con la empresa de estudio. Las informaciones generadas permiten realizar el marco comparativo entre la ejecución real de cada una de sus variables y el valor meta (eficiencia en tiempos de proceso por año, *batch time* por año, eficiencia en tiempos por mes, *batch time* mes, eficiencia en tiempos por día, *batch time* diario).

**Figura 3.** Ejemplo de Indicadores de desempeño KPI



Fuente: Elaboración propia (2021)

### III. Resultados

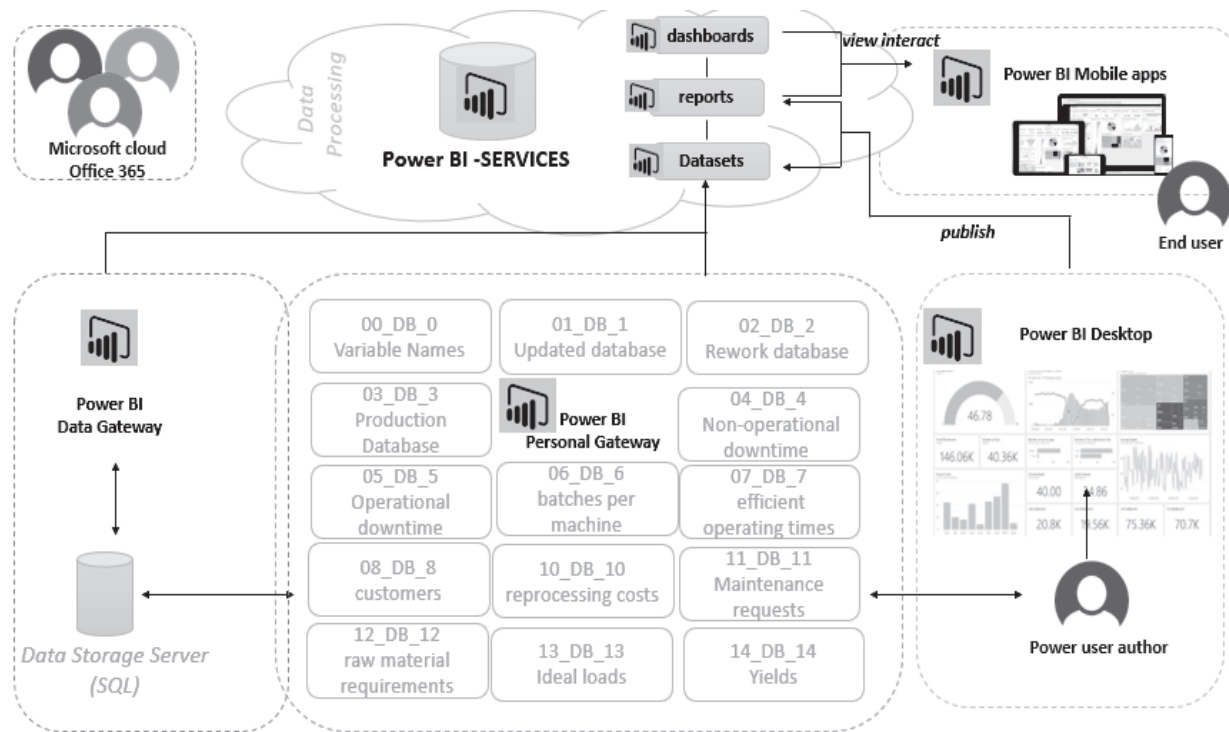
El comportamiento industrial en el caso de estudio ha permitido desarrollar un nuevo e innovador concepto de integración del conjunto de informaciones en tiempo real, que permite responder a las múltiples necesidades desde cada una de las áreas operativas: calidad, administración, gestión de proyectos, logística y transporte, entre otras. Esto se logra con la correcta aplicación de la tecnología Big Data y, de acuerdo con [20] y [21], este tipo de innovación permite un mayor y mejor control sobre los procesos organizacionales de la compañía, así como la simplificación y el aumento en la eficacia del desempeño del conjunto de informaciones relacionadas, todo lo anterior en pro de la correcta toma de decisiones organizacionales, fundamentadas en la validez y eficacia informativa en tiempo real.

Para conseguir los resultados expresados anteriormente, fueron generados los principales tableros de visualización y control a partir del conjunto de informaciones (bases de datos) de la empresa, así como el correcto análisis y seguimiento de los procesos operativos, presentados a continuación.

### 6. Arquitectura de datos en el modelamiento Big Data: caso de estudio textil

La arquitectura de datos representa, de acuerdo con [22] y [23], el modelamiento, políticas y sistemas estandarizados de recopilación de las informaciones, así como su enrutamiento en el comportamiento operacional. A la vez, permite identificar en tiempo real el marco de integración del conjunto de informaciones, así como sus niveles de asociatividad entre departamentos (ver Figura 4).

**Figura 4.** Arquitectura de datos textil: caso de estudio



Fuente: Elaboración propia (2021)

El modelo de arquitectura de datos para la presente investigación en la industria textil es presentado en la Figura 4. La integración de conjunto de datos está representada por dos módulos Power BI para su correcta sincronía: Power BI Data Gateway, siendo el enlace y puerta de acceso al sistema de información de la empresa, donde se procesa y almacena en tiempo real el conjunto de informaciones de los procesos internos y externos (Data Storage Server - SQL). Una vez son extraídos del sistema SQL, los datos son distribuidos de acuerdo con su origen, a cada base de datos asignada para su correcto almacenamiento en cada departamento de la empresa. Lo anterior es posible gracias a Power BI Personal Gateway y a la importación de los mismos al conjunto de bases de datos de la compañía. Con el conjunto de informaciones disponibles en la nube de almacenamiento de la organización, se procede a realizar el procedimiento descrito en la presente investigación, posible con el módulo Power BI

Services: Dataset, para el procesamiento, depuración y limpieza del conjunto de datos requeridos, *reports*, en la generación del análisis de las variables deseadas y la generación de los *dashboard* o tableros de visualización, donde se representan gráficamente los resultados obtenidos de la interacción y generación multiescenarios de los procesos operativos, así como el análisis de las variables de estudio.

Como resultado de salida en el procesamiento de las informaciones, Power BI Desktop se presenta como una excelente herramienta de software para la visualización, seguimiento y control operativo interno gerenciado por operarios, jefes de planta y coordinadores. Para el análisis desde el departamento de gestión estratégica y visualización de informes en escenarios externos, es utilizado Power BI Mobile Apps, permitiendo, en tiempo real, la visualización de informes desde los dispositivos móviles y el desarrollo de las necesidades informativas.

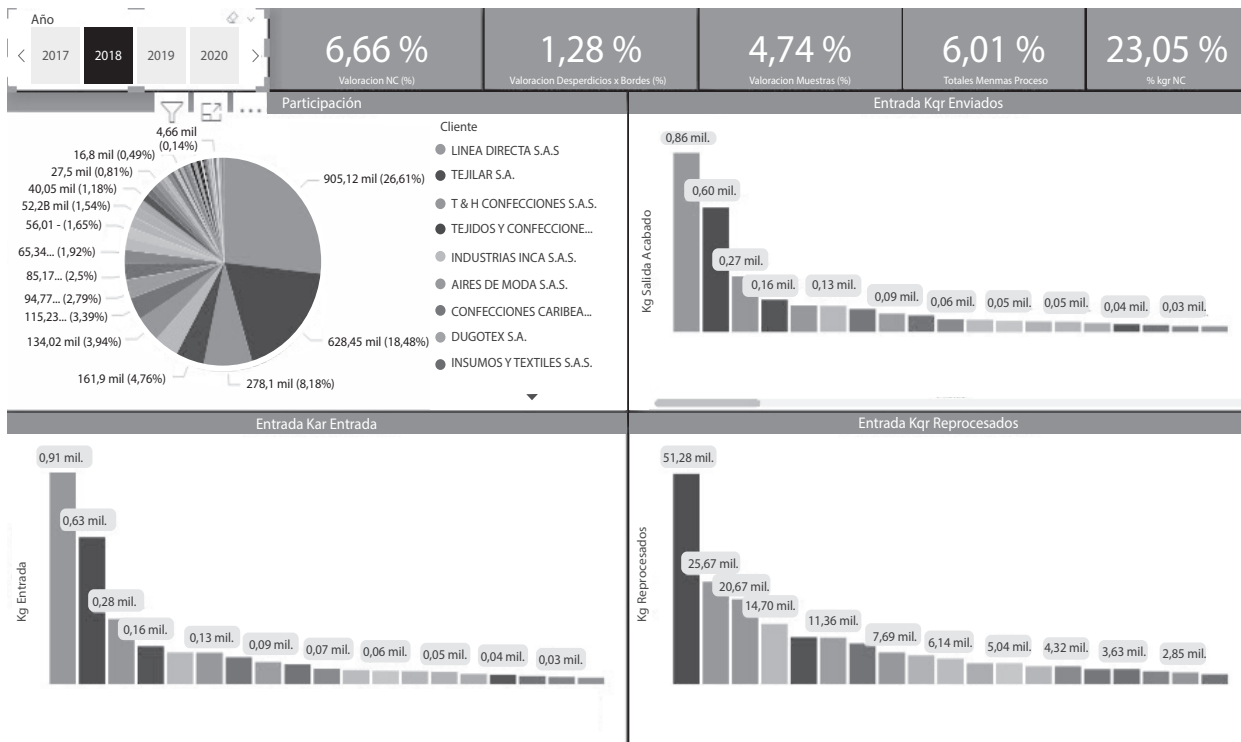
### 7. Análisis operativo interno textil: niveles de procesamiento de kilogramos

Algunos de los elementos clave en los estudios relacionados con la ciencia de los datos son la correcta aplicación de las informaciones, su colocación en tiempo real y su uso en la toma de decisiones gerenciales. Para ello existen múltiples herramientas de visualización, al interior del software Power BI, que permiten la correcta planeación, visualización y analítica de datos en tiempo real. Este conjunto e integración de elementos visuales al interior de un entorno de visualización se conoce como tablero de visualización o *dashboard*.

En el caso de la presente investigación del sector textil, se crea un tablero de visualización para el análisis de las operaciones (ver Figura 5). En este tipo de tablero visual puede observarse:

porcentaje de participación de los clientes, en el cual se relaciona el margen de contribución en ventas de cada uno de los consumidores de servicios de la empresa de estudio; *entrada de kg enviados*, referentes a los kilogramos de producto (tela) enviados a la empresa para su posterior procesamiento. Debe tenerse en cuenta que no necesariamente todos los kilogramos que se envían ingresan al sistema productivo, debido a que se realiza una inspección previa de la calidad del insumo entrante (tela); este valor neto de entrada de los insumos, una vez realizada la inspección de calidad, se conoce como la variable *entrada kg de entrada*. Finalmente, es analizada en el sistema de operaciones de la compañía la *entrada de kg reprocesados*, indicador que define la cantidad de kilogramos que fueron reprocesados al interior del sistema productivo, una vez realizada la transformación de los insumos en la operación industrial de teñido.

Figura 5. Tablero de visualización del análisis de operaciones de estudio.



Fuente: Elaboración propia (2021)



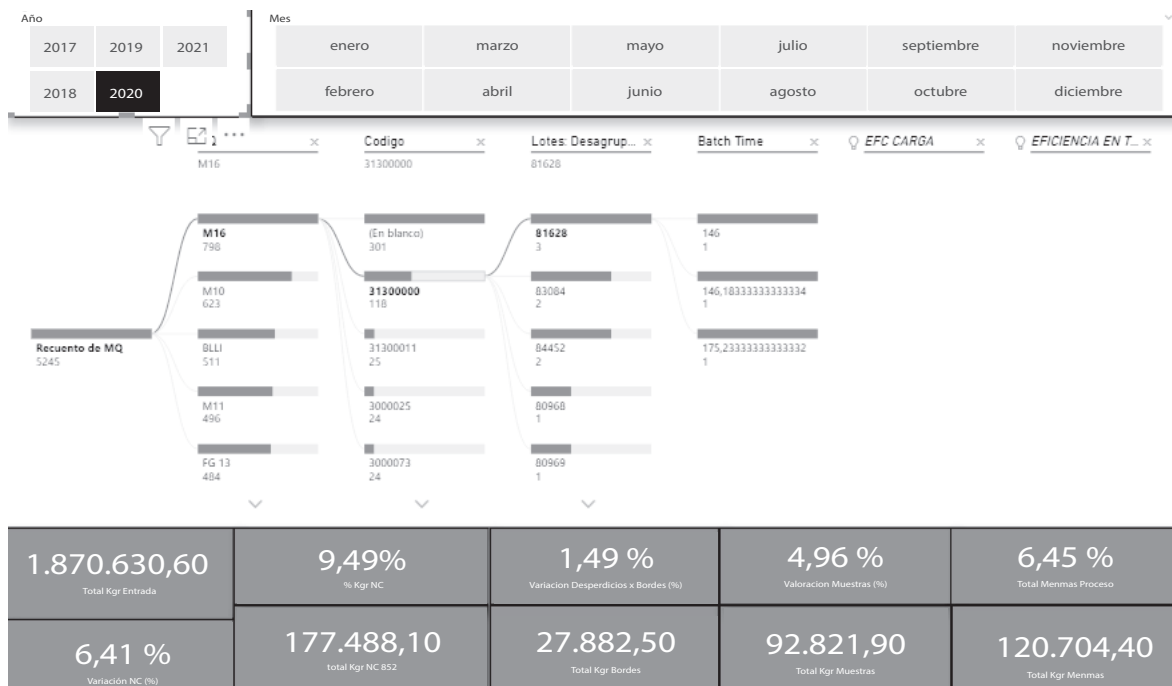
En la analítica de datos de la visualización expuesta, es importante destacar que la diferencia entre *entrada kg de entrada* y *entrada de kg reprocesados* genera como resultado el valor del nivel de procesamiento de kilogramos, cifra de valor neto de los insumos procesados de forma correcta al interior del sistema productivo. Como puede observarse, existen ciertas referencias de clientes que consumen servicio de procesamiento en bajas cantidades; sin embargo, sus niveles de reprocesamiento industrial son elevados, lo que indica que el tipo de servicio y la forma secuencial de los procesos a través de los cuales es solicitado el producto final afectan el sistema operativo industrial de la empresa, generando una relación costo-beneficio negativa en función de los objetivos de rentabilidad organizacionales.

### 8. Análisis en el procesamiento y control de los lotes de producción

En la analítica avanzada del estudio y ciencia de los datos, existen aplicaciones fundamentales

para la resolución de problemas de índole inferencial en función de los niveles de especificidad que pueden lograrse con las informaciones. En relación con lo anterior, la inteligencia artificial (IA) existe como una alternativa de solución, la cual es definida por [24] como la integración de algoritmos inteligentes capaces de analizar grandes cantidades de datos en el procesamiento rápido e iterativo, con el propósito de identificar de forma sistematizada características de comportamiento y patrones en los datos. Para la investigación desarrollada, un tipo de técnica perteneciente a IA aplicada es conocida como *data mining* o minería de datos, la cual apoya el proceso de exploración y análisis de grandes conjuntos de datos, en búsqueda de comportamientos o patrones a ser analizados [25], [26]. Para este caso de estudio, se realizó una tipología de *data mining* al interior del sistema operativo de las máquinas industriales, como se expone en la Figura 6.

Figura 6. Tablero de visualización (IA) de análisis operacional de maquinaria textil.



Fuente: Elaboración propia (2021)

Una de las principales aplicaciones desarrolladas para el control de análisis de las operaciones de las máquinas industriales, utilizando el software Power BI, es conocida como gráfico de esquema jerárquico. La gran ventaja que ofrece este elemento (visualización de IA) es la exploración a profundidad del conjunto de variables que requieren ser analizadas en la programación de los equipos industriales de la compañía. En la Figura 6 se presenta un ejemplo aplicado para la máquina M16, cuya actividad fundamental es el teñido de tela a tonalidades claras y oscuras. Como puede observarse en la parte central del gráfico de esquema jerárquico, una vez se identifica el número de lote (5245) en el clúster de maquinaria *recuento MQ*, el modelo de datos identifica automáticamente, a través de exploración inferencial de información, la máquina (M16) en la cual se realizó este tipo de lote; a continuación, el sistema identifica el código, el cual es la asignación de tipo de procesamiento de teñido aplicado por la máquina; en la siguiente fase del modelo, el lote mayor (5245) es desagrupado en diversos tipos de lotes desagrupados y, en el presente caso, para este referente de máquina fue asignado el acondicionamiento operativo del lote desagrupado (81628), considerando el comportamiento que se genera, debido a que en cada tratamiento de tonalidad afectan ciertos tipos de variables como temperatura, tonalidad de químico, tiempo de procesamiento, proceso de secado, entre otras, y, por ende, debe ser asignado un tipo de máquina específico para realizar este tipo de proceso. Finalmente, se presenta el *batch time* o tiempo de lote, que es el tiempo de procesamiento desde la entrada del insumo a la máquina hasta que se entrega el lote procesado (tiempo total). Como puede observarse, existen tres requerimientos de entrada y procesamiento de lote (*batch time*), cada uno con un tiempo de procesamiento definido según la cantidad de kilogramos procesados en la máquina asignada.

Con el propósito de llevar a cabo un mayor y mejor control de las actividades operativas de cada una de las máquinas, fueron generados indicadores claves de desempeño (KPI), claves para la medición del rendimiento de cada uno de los equipos industriales de estudio, el cual se ilustra en la

parte inferior de la Figura 6 a través de tarjetas de medición. Como puede observarse, con la identificación de análisis de la máquina 16 (M16), se activan automáticamente el total de kilogramos de entrada (1.870.630), el porcentaje de kilogramos no conformes (9,49%), el porcentaje de variación de desperdicios por bordes (1,49%), el porcentaje de variación de muestras (4,96%), el porcentaje del total de mermas del proceso (6,45%), el porcentaje de variación de producto no conforme (6,41%), el total de kilogramos no conformes (177.488), el total de kilogramos de borde (27.882), el total de kilogramos de muestra (92.821) y el total de kilogramos de mermas (120.704). Estos resultados pueden ser monitoreados en tiempo real y, de forma automática, realizar un análisis inferencial en la identificación de las muestras de lotes desagrupados con mayor número de no conformes en sistema, todo esto en pro de mejorar la eficiencia y trazabilidad del proceso de teñido de los servicios requeridos por los clientes en la empresa de estudio.

#### IV. Discusión

El conjunto de aplicaciones generadas a partir de la tecnología Big Data para la presente investigación impulsa la correcta planeación, implementación, control y seguimiento de los procesos operacionales de estudio a través del análisis de datos, así como la generación de planes de mejora en tiempo real en relación con la analítica del ecosistema industrial textil de estudio: operaciones de maquinaria, sistemas de información, recurso humano, recurso financiero, y es precisamente la sinergia que se genera entre los elementos anteriores, a través de la conexión de los sistemas de información en el modelo de datos, lo que permite generar, como resultado, recursos visuales a través de tableros de visualización (*dashboard*) que permiten un monitoreo en tiempo real y la subsecuente toma de decisiones empresariales.

Los datos analizados en el presente estudio son el resultado de las bases de datos donde se presenta la información relacionada para cada lote de pedido de las principales variables a ser analizadas: cantidad de kilogramos de productos

de entrada, kilogramos procesados, kilogramos no conformes, kilogramos de salida, entre otras. Es importante resaltar que la analítica de datos se enfoca en el número y tipo de pedidos que realiza cada uno de los clientes a la organización, ya que la forma y presentación de los requerimientos de producción afecta directamente la programación de maquinaria y equipos internos, así como todo un proceso logístico que demanda cada pedido. Por ello, resulta fundamental la identificación de

cada colocación de lote de pedido para cada tipo de cliente y la trazabilidad que conlleva este tipo de proceso en el horizonte de tiempo, ya que el impacto en costos asociados influye directamente en el margen de utilidad de la empresa del sector textil de estudio. Con el fin de generar un mayor y mejor control operacional de los pedidos y su comportamiento en periodos de tiempo, se creó un tablero de visualización del índice de participación, presentado en la Figura 7.

**Figura 7.** Tablero de visualización del índice de participación operacional de pedidos.



Fuente: Elaboración propia (2021)

Un estudio del comportamiento trimestral de los kilogramos procesados y kilogramos no conformes es presentado en las tablas de la Figura 7. Un ejemplo del comportamiento o patrón de este conjunto de datos se presenta a continuación: Puede observarse que, para el segundo trimestre del año 2017, se generó una producción de 135.057 kg; si se compara con el dato de producción del tercer trimestre del mismo año (867.988 kg), puede observarse que se presenta un incremento porcentual de 6,17%, lo cual permite identificar, a través de este tipo de análisis, los márgenes de incremento o reducción de la producción en los periodos definidos. Ahora, si se hace un análisis paralelo con la tabla de kilogramos no conformes (derecha), puede observarse, por ejemplo, que en el segundo y tercer trimestres del año 2017 se generaron 35.648 kg y 328.793 kg no conformes, respectivamente. Como se ve, a cada tasa de producción en kilogramos por cada trimestre, se genera cierta cantidad de producto que debe ser procesada, y este tipo de comportamiento es uno de los propósitos de la investigación: realizar el análisis inferencial, a partir de esta cantidad de kilogramos reprocesados o no conformes, por ejemplo, del tercer trimestre (328.793) de cuál es el equivalente de acuerdo con cada lote de pedido de clientes y cómo está impactando en los reprocesos internos que tiene la organización. Para ello se plantea como propuesta la transformación de un escenario de responsabilidad única, donde la compañía debe responder por la totalidad de sus reprocesos, a un escenario de responsabilidad compartida, donde, una vez sean analizados todos y cada uno de los clientes, se les pueda dar una inducción de la forma correcta de programar sus pedidos y evitar la afectación operacional interna que conlleva el tipo de programaciones que generan.

Un segundo tipo de análisis puede observarse en función del porcentaje de variación de la cantidad de kilogramos procesados, realizando el comparativo entre los dos periodos de análisis, como puede observarse en los gráficos de barras centrales. Un ejemplo del comportamiento o patrón de este conjunto de datos se presenta a continuación: la cantidad de kilogramos procesados en el segundo trimestre del año 2017 es de 135.057 kg y la cantidad de

kilogramos procesados en el segundo trimestre del año 2018 es de 809.222 kg, presentando una tasa de incremento de 499,17%. El análisis anterior es generado también para la variable *kilogramos no conformes*. Finalmente, se evidencia, a través del presente estudio, que este tipo de análisis permite que la compañía pueda identificar el comportamiento operacional de cada periodo y controlar la aleatoriedad de la producción, mejorando la programación de operaciones y gestión de recurso financiero y físico. Los resultados del presente estudio permitieron evidenciar una mejora del 20% en la identificación de fallas operacionales y no operacionales del proceso industrial interno.

## V. Conclusiones

La implementación de la tecnología Big Data al interior de los procesos operacionales en el sector textil de estudio permite generar una sinergia e integración de las informaciones de las diferentes áreas: industrial, financiera, calidad y gestión administrativa. El uso eficiente de datos se logra a través del correcto desarrollo metodológico, enfocado en ciencia, de los datos, en cada una de sus fases: obtención de insumos de bases de datos, depuración y limpieza, generación de modelo de datos, visualización y análisis, proyección y comunicación de informes.

El software Power BI se presenta como una eficiente herramienta de inteligencia de negocios, a través de la cual se aplica tecnología Big Data para el desarrollo del caso de estudio. El uso y gerenciamiento masivo y la estandarización de los datos permiten una mejor planeación, control y seguimiento de los procesos organizacionales a través de la generación de indicadores clave de desempeño, analítica de series de tiempo de producción, plan de acción y oportunidades de mejora ante producto no conforme y paros de maquinaria. La creación de los tableros de visualización, por su parte, permite generar la transformación de un proceso de gerenciamiento operacional de forma "empírica" a un escenario tecnológico, obteniendo una mejora del 20% en la identificación de fallas propuestas, y de mejora de la productividad en un 15%.

De acuerdo con los resultados obtenidos en la presente investigación, es recomendable realizar un estudio específico de clientes en relación con las categorías de tipo, cantidad y fechas de colocación de pedidos en la programación industrial. La correcta capacitación y formación de los proveedores en temáticas como programación de pedidos, características de operación, poder de negociación de pedidos y estudio de productividad interna y externa se presenta como un requerimiento a tener en cuenta, todo lo anterior en el marco de una filosofía de responsabilidad productiva que la empresa debe implementar para conseguir sus metas de productividad y rentabilidad a mediano y largo plazo.

## VI. Referencias

- [1] M. Chen, S. Mao y Y. Liu, "Big Data: A Survey", *Mobile Networks and Applications*, vol. 19, no. 2, pp. 171-209, 2014 [En línea]. Disponible en: <https://doi.org/10.1007/s11036-013-0489-0>
- [2] C. J. Cremin, S. Dash y X. Huang, "Big data: Historic advances and emerging trends in biomedical research", *Current Research in Biotechnology*, vol. 4, pp. 138-151, enero de 2022 [En línea]. Disponible en: <https://doi.org/10.1016/j.crbiot.2022.02.004>
- [3] D. Wiltshire y S. Alvanides, "Ensuring the ethical use of big data: lessons from secure data access", *Heliyon*, vol. 8, no. 2, febrero de 2022 [En línea]. Disponible en: <https://doi.org/10.1016/j.heliyon.2022.e08981>
- [4] Y. Zhang y J. Li, "Performance analysis of big data transmission in wearable system based on special textile clothing", *Procedia Computer Science*, vol. 183, pp. 713-719, 2021 [En línea]. Disponible en: <https://doi.org/10.1016/j.procs.2021.02.119>
- [5] J. Yang, Y. Li, Q. Liu, L. Li, A. Feng, T. Wang, S. Zheng, A. Xu y J. Lyu, "Brief introduction of medical database and data mining technology in big data era", *Journal of Evidence-Based Medicine*, vol. 13, no. 1, pp. 57-69, febrero de 2020 [En línea]. Disponible en: <https://doi.org/10.1111/jebm.12373>
- [6] C. A. Escobar, M. E. McGovern y R. Morales-Menendez, "Quality 4.0: a review of big data challenges in manufacturing", *Journal of Intelligent Manufacturing*, vol. 32, no. 8, pp. 2319-2334, 2021 [En línea]. Disponible en: <https://doi.org/10.1007/s10845-021-01765-4>
- [7] G. Raviv, *Collect, Combine, and Transform Data Using Power Query in Excel and Power BI*. Microsoft Press, 2018.
- [8] B. Powell, *Microsoft Power BI Cookbook: Creating Business Intelligence Solutions of Analytical Data Models, Reports, and Dashboards*. Packt Publishing, 2017.
- [9] A. Aspin, *Pro Power BI Desktop*. Apress Berkeley, 2018 [En línea]. Disponible en: <https://doi.org/10.1007/978-1-4842-3210-1>
- [10] L. T. Becker, colab., y E. M. Gould, ed., "Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data", *Serials Review*, vol. 45, no. 3, pp. 184-188, 2019 [En línea]. Disponible en: <https://doi.org/10.1080/00987913.2019.1644891>
- [11] M. Cornejo, F. Mendoza y R. C. Rojas, "La Investigación con Relatos de Vida: Pistas y Opciones del Diseño Metodológico", *Psykhé*, vol. 17, no. 1, pp. 29-39, 2008 [En línea]. Disponible en: <https://doi.org/10.4067/S0718-22282008000100004>
- [12] A. C. Salgado Lévano, "Investigación cualitativa: diseños, evaluación del rigor metodológico y retos", *Liberabit*, vol. 13, no. 3, pp. 71-78, 2007.
- [13] J. A. Yuni y C. A. Urbano, *Técnicas para investigar: recursos metodológicos para la preparación de proyectos de investigación*, 2ª ed. Córdoba: Brujas, 2006.
- [14] W. J. Sotaquirá Ayala, "Power BI como herramienta de Big Data & Business Analytics para Onelink Colombia", Tesis de maestría. 2017. Repositorio Institucional Universidad EAFIT [En línea]. Disponible en: <http://hdl.handle.net/10784/11767>

- [15] N. Cohen y G. Gómez Rojas, *Metodología de la investigación, ¿para qué?: la producción de los datos y los diseños*. Buenos Aires: Teseo, 2019 [En línea]. Disponible en: <http://up-rid2.up.ac.pa:8080/xmlui/handle/123456789/1363>
- [16] O. A. Agudelo-Giraldo, J. E. León-Molina, M. A. Prieto-Salas, A. Alarcón-Peña y J. C. Jiménez-Triana, *La pregunta por el método: derecho y metodología de la investigación*. Bogotá: Universidad Católica de Colombia, 2018 [En línea]. Disponible en: <https://hdl.handle.net/10983/22541>
- [17] K.-J. Villegas-Sandoval F. Martínez-Olmo, "Tendencias en el diseño metodológico de investigación sobre la evaluación de competencias en la educación superior", *REIRE. Revista d'Innovació i Recerca en Educació*, vol. 10, no. 1, pp. 1-13, 2017 [En línea]. Disponible en: <http://doi.org/10.1344/reire2017.10.11011>
- [18] I. Ferraz Pinto, C. J. Gomes Campos y C. Siqueira, "Investigação qualitativa: Perspetiva geral e importância para as Ciências da Nutrição", *Acta Portuguesa de Nutrição*, no. 14, pp. 30-34, 2018 [En línea]. Disponible en: <https://dx.doi.org/10.21011/apn.2018.1406>
- [19] O. A. Camacho Hernández, "Propuesta de implementación de una solución de inteligencia de negocios para el Área de Gestión Tributaria de la Municipalidad de El Guarco", Trabajo de grado. Cartago, 2021. Repositorio TEC [En línea]. Disponible en: <https://hdl.handle.net/2238/13708>
- [20] C.-W. Tsai, C.-F. Lai, H.-C. Chao y A. V. Vasilakos, "Big data analytics: a survey", *Journal of Big Data*, vol. 2, 2015 [En línea]. Disponible en: <https://doi.org/10.1186/s40537-015-0030-3>
- [21] S. Kumar, D. Sharma, S. Rao, W. M. Lim y S. K. Mangla, "Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research", *Annals of Operations Research*, 2022 [En línea]. Disponible en: <https://doi.org/10.1007/s10479-021-04410-8>
- [22] B. Bansal, V. N. Jenipher, R. Jain, R. Dilip, M. Kumbhkar, S. Pramanik, S. Roy y A. Gupta, "Big Data Architecture for Network Security", en *Cyber Security and Network Security*, S. Pramanik, D. Samanta, M. Vinay y A. Guha, eds., pp. 233-267. Scrivener Publishing, 2022 [En línea]. Disponible en: <https://doi.org/10.1002/9781119812555.ch11>
- [23] E. N. Witanto, Y. E. Oktian y S.-G. Lee, "Toward Data Integrity Architecture for Cloud-Based AI Systems", *Symmetry*, vol. 14, no. 2, 2022 [En línea]. Disponible en: <https://doi.org/10.3390/sym14020273>
- [24] C. Debrah, A. P. C. Chan y A. Darko, "Artificial intelligence in green building", *Automation in Construction*, vol. 137, mayo de 2022 [En línea]. Disponible en: <https://doi.org/10.1016/j.autcon.2022.104192>
- [25] R. Safdari, A. Deghatipour, M. Gholamzadeh y K. Maghooli, "Applying data mining techniques to classify patients with suspected hepatitis C virus infection", *Intelligent Medicine*, vol. 2, no. 4, pp. 193-198, noviembre de 2022 [En línea]. Disponible en: <https://doi.org/10.1016/j.imed.2021.12.003>
- [26] I. Batool y T. A. Khan, "Software fault prediction using data mining, machine learning and deep learning techniques: A systematic literature review", *Computers and Electrical Engineering*, vol. 100, mayo de 2022 [En línea]. Disponible en: <https://doi.org/10.1016/j.compeleceng.2022.107886>