



OPTIMIZACIÓN ENERGÉTICA EN EL PROCESO DE SECADO DE PAPEL TISSUE, APLICANDO ALGORITMOS DE MACHINE LEARNING, BUSCANDO EL MÍNIMO COSTO

**ENERGY OPTIMIZATION IN THE TISSUE PAPER DRYING PROCESS, APPLYING MACHINE LEARNING ALGORITHMS,
SEEKING THE MINIMUM COST**

Jesús Alejandro Moreno Salas¹

Politécnico Colombiano Jaime Isaza Cadavid, PCJIC

Henry Omar Sarmiento Maldonado²

Politécnico Colombiano Jaime Isaza Cadavid, PCJIC

Jorge Alberto Londoño Pulgarín³

Politécnico Colombiano Jaime Isaza Cadavid, PCJIC

RESUMEN

El objetivo de este trabajo consiste en determinar las variables que más afectan al proceso industrial de secado de papel y sus rangos de operación para encontrar la combinación óptima del menor costo en el consumo de electricidad, gas y vapor. La utilización de Machine Learning, específicamente Redes Neuronales Artificiales y el método estadístico de Regresión Lineal Multivariada con algoritmos aplicados exitosamente a nivel industrial, permitió extraer información del historial de datos y, a partir de allí, encontrar los valores óptimos de operación que hacen posible un ahorro energético del 5.18% en la producción de papel Tissue.

PALABRAS CLAVE: aprendizaje automático; consumo de energía primaria; industria papelera; papel higiénico; secado de papel.

ABSTRACT

The objective of this work is to determine the variables that most affect the industrial paper drying process and their operating ranges to find the optimal combination of electricity, gas, and steam consumption with the lowest cost. The use of Machine Learning, specifically Artificial Neural Networks, and the statistical method of Multivariate Linear Regression with algorithms successfully applied at the industrial level, made it possible to extract information from historical data and, from there, find the optimal operating values that enable energy savings of 5.18% in tissue paper production.

KEYWORDS: machine learning; paper drying; paper industry; primary energy consumption; toilet paper.

¹ Jesús-Alejandro Moreno-Salas [CV] Aspirante al título de Magister en ingeniería de gestión de la automatización - Politécnico Colombiano Jaime Isaza Cadavid, PCJIC, Colombia - alejandro.moreno13669@gmail.com

² Henry-Omar Sarmiento-Maldonado [CV] PhD en Ing. Electrónica, Profesor Asociado. (Facultad de Ingenierías) - Politécnico Colombiano Jaime Isaza Cadavid, PCJIC, Colombia - hosarmiento@elpoli.edu.co

³ Jorge-Alberto Londoño-Pulgarín [CV] Magister en Automatización, Profesor de catedra (Facultad de Ingenierías)- Politécnico Colombiano Jaime Isaza Cadavid, PCJIC, Colombia - jorgelondono@elpoli.edu.co

RESUMO

O objetivo deste trabalho é determinar as variáveis que mais afetam o processo de secagem industrial de papel e suas faixas de operação, a fim de encontrar a combinação ideal de consumo de eletricidade, gás e vapor com o menor custo. O uso de Aprendizagem de Máquina, especificamente Redes Neurais Artificiais, e o método estatístico de Regressão Linear Multivariada com algoritmos aplicados com sucesso em nível industrial, permitiram extrair informações de dados históricos e, a partir daí, encontrar os valores operacionais ótimos que possibilitam uma economia de energia de 5,18% na produção de papel tissue.

PALAVRAS-CHAVE: aprendizagem automática; consumo de energia primária; indústria do papel; papel higiênico; secagem de pape.

1. INTRODUCCIÓN

En procesos industriales el crecimiento, debido a la demanda de productos, ha hecho que se utilicen sistemas de control e instrumentación con el fin de mantener los procesos estables, eficientes y llevándolos a disminuir costos. Sin embargo, los procesos no son completamente automáticos y es usual que haya variables que los operarios deben manipular constantemente, sea de forma digital o incluso de forma manual basados en su conocimiento o experiencia; práctica que conlleva a posibles fallos o deterioro de las condiciones óptimas de producción.

Ya desde hace varias décadas, pero actualmente con mayor facilidad y menor costo, las variables de interés para el control de los sistemas automáticos se almacenan en bases de datos (data set), recolectándose así la información histórica del proceso que normalmente es utilizada como fuente de verificación de situaciones o eventos ocurridos para la mejora en línea de los mismos.

En la última década la industria de procesos ha venido reformulándose y en Alemania se propuso un nuevo concepto denominado Industria 4.0. Esta nueva tendencia tiene como objetivo principal desarrollar industrias inteligentes. En este sentido, las empresas informáticas tan importantes en la actualidad se han venido concentrando en recolectar, procesar, almacenar y distribuir información que pueda ser de gran valor y utilidad para diferentes propósitos, entre otros, técnicos, económicos y gerenciales. Como consecuencia, el análisis de información a partir del historial de datos ofrece información valiosa la cual genera un conocimiento nuevo que facilita la comprensión y toma de decisiones dentro de procesos industriales (Witten et al., 2016).

Desde hace varios años, los procesos de manufactura en las empresas de fabricación de papel cuentan con equipos especializados que están instrumentados y realizan los controles correspondientes para garantizar procesos limpios y sistematizados. Estas empresas tienen los sistemas de producción monitoreados por instrumentos ubicados a lo largo de sus procesos, los cuales generan información que es acumulada en servidores. El análisis de estos datos juega un papel importante para el descubrimiento de información valiosa con la cual se puede tomar decisiones de mejora (Maisueche Cuadrado, 2019).

También, durante las últimas décadas los sistemas computacionales han sido utilizados para analizar datos mediante herramientas que extraen información, desde la perspectiva del aprendizaje automático, conocido en inglés como “Machine Learning (ML)”. La inteligencia artificial como herramienta del ML emplea algoritmos que ofrecen una autonomía para optimización en los diferentes procesos donde se aplica. Uno de los objetivos más aplicados de esta herramienta es maximizar la eficiencia identificando patrones complejos a partir de grandes volúmenes de datos, procesándolos para predecir comportamientos (Aldrich & Auret, 2013).

El presente trabajo busca analizar datos históricos registrados y almacenados en una empresa del sector de manufactura del papel higiénico en el proceso de secado de papel “Tissue”, trabajo direccionado a obtener información importante en materia de gestión energética. Se busca obtener conocimiento desde los datos validando el conocimiento que tiene la empresa desde la experiencia para definir los rangos óptimos del proceso. Esto permite establecer qué variables hacen que el proceso sea más eficiente y en qué rangos deben estar para reducir los costos de energía eléctrica, costo de gas y costo de vapor.

Como cifras de referencia que justifican el objetivo planteado, se destacan los valores asociados al costo final de la producción en la industria del papel donde el 8% corresponde al consumo de energía, el 65% al costo de la materia prima y el 27% a costos administrativos (Caluña & Patricio, 2018). Además, asociada a esta industria se estima que su requerimiento es de aproximadamente el 6% del total de la demanda de energía industrial en todo el mundo. En esta industria con un consumo significativo de energía, la sección de secado es aproximadamente el 50% del consumo del total en una fábrica de papel (Laurijssen et al., 2010).

En razón de lo expuesto anteriormente, este proyecto busca avanzar en la optimización de variables energéticas asociadas al proceso de secado de papel higiénico de doble hoja, con el uso de ML. Iniciando con la presentación de la metodología del trabajo, se incluye la descripción del proceso destacando los diferentes elementos de relevante importancia para la propuesta, donde, desde la experiencia de los expertos del

proceso, se realiza análisis profundo de las variables a optimizar mediante métodos estadísticos y algoritmos inteligentes que entreguen una mejora cuantificable y además genere un modelo para predecir costos de manufactura rápidos.

2. MÉTODO

Para el presente trabajo se adoptó una metodología dividida en fases, las cuales permitieron una ejecución de tareas ordenada y estructurada, las fases fueron:

Fase 1. Investigación.

Fase 2. Recolección de datos.

Fase 3. Preprocesamiento de datos y generar Data Set.

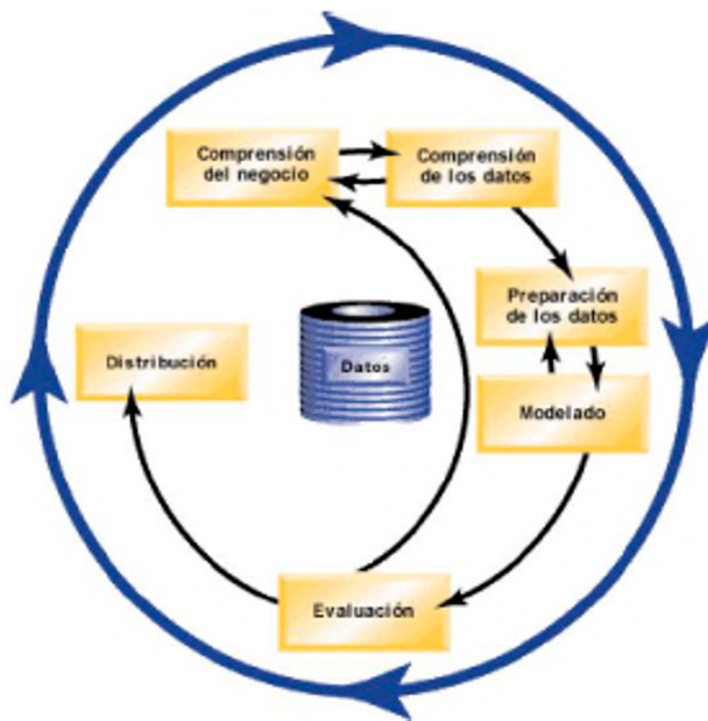
Fase 4. Entrenamiento de algoritmos.

Esta metodología busca realizar un estudio detallado del funcionamiento de la planta, pasando por todos sus procesos, con el objetivo principal de evidenciar la crucial importancia del proceso de secado de papel e identificar de forma general el proceso de fabricación completo. Luego, con la información recolectada por los expertos, identificar los valores y variables de mayor influencia en el costo operativo asociados a la producción del papel.

Para el análisis de datos se realizó el preprocesamiento que permitió tener datos limpios y confiables y con estos hacer los entrenamientos. Se usó el método estadístico de Regresión Lineal Multivariable y el algoritmo de Redes Neuronales Artificiales para generar un modelo confiable de estimación del costo final del producto.

Se usó como referencia la metodología CRISP-DM (Cross Industry Standard Process for Data Mining), metodología utilizada y orientada para trabajos de minería de datos, con 6 etapas que pueden ser bidireccionales. Es decir, si en una etapa no se satisface la necesidad es necesario devolverse a la anterior. En la Figura 7 se ve la secuencia entre las etapas.

Figura 7 . Etapas de la implementación de la metodología CRISP-DM



Fuente: (Marques, 2014)

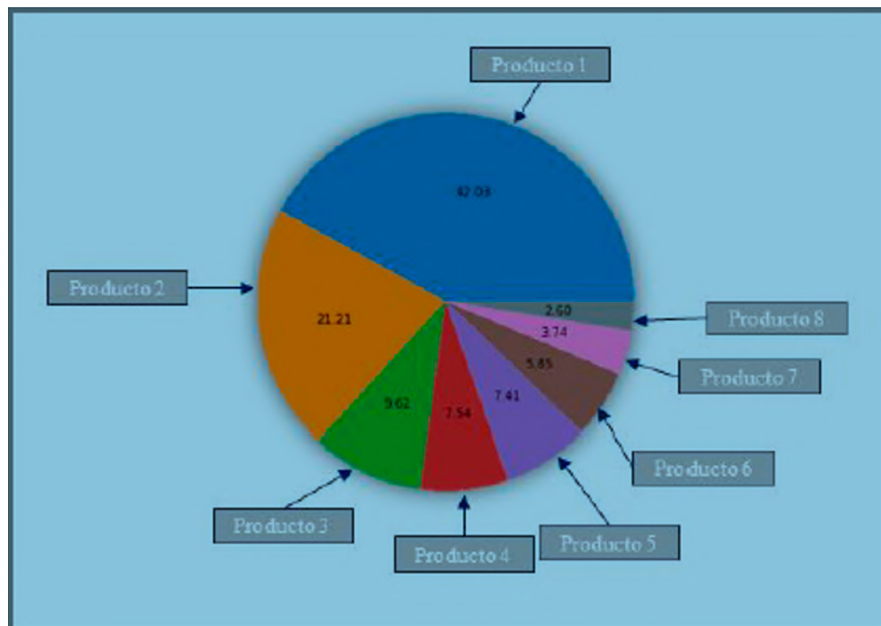
2.1. INVESTIGACIÓN

En esta fase se realiza toda la comprensión del proceso, desde lo general hasta lo específico, en la fabricación de papel Tissue. Se identificaron 3 grandes procesos principales: la preparación de pasta, la manufactura de grandes rollos y el doblado y/o corte para convertir en producto final. En la Figura 10 se puede apreciar todo el proceso productivo en un diagrama de flujo.

El enfoque de este trabajo está en la manufactura de grandes rollos (ver Figura 9) específicamente en el proceso de secado de papel por calentamiento, que es el proceso final en la producción de grandes rollos.

El papel Tissue es un tipo de papel sanitario con características de suavidad, elasticidad y absorción, producido para hacer papel higiénico, toallas, servilletas, pañuelos, rollos de cocina, faciales, sabanillas para clínicas y hospitales y papeles para dispensadores. La fabricación de este papel se hace con materia pura (pulpa virgen) y con papel reciclado en concentraciones definidas por la empresa. El objeto de estudio de este trabajo se centra solo en un producto el de mayor producción dentro de la industria el cual es el papel sanitario higienico que representa 42 % de la manufactura dentro de 8 productos que se fabrican en esta fábrica ver Figura 8.

Figura 8. Productos de fabricación y su porcentaje de manufactura dentro de la fábrica



Fuente: Completamente generada por el autor, 2023

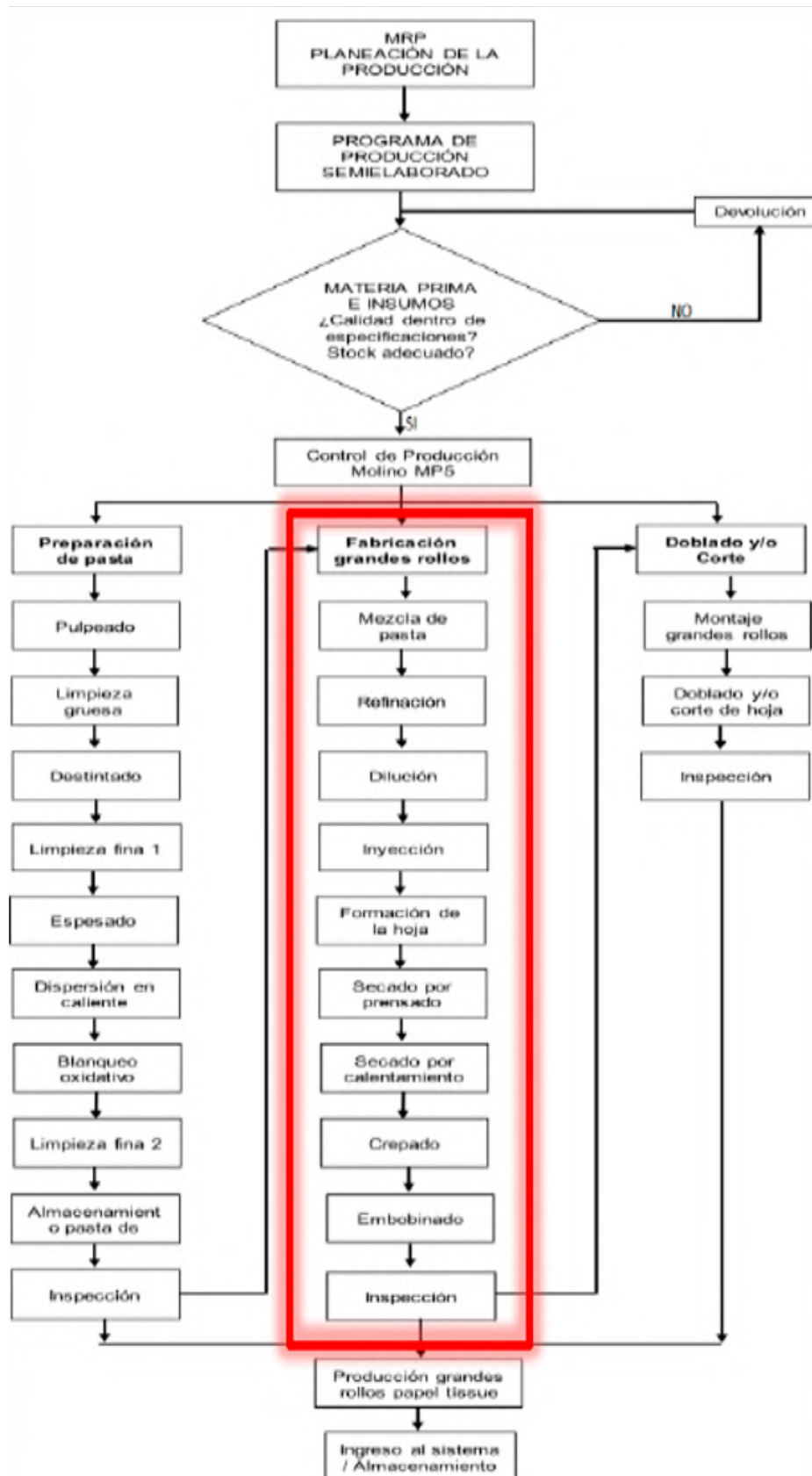
En el apartado final se realiza la investigación y análisis de los diferentes algoritmos y métodos estadísticos más usados en el aprendizaje automático conocido en inglés como Machine Learning (ML), con el objetivo de garantizar el rendimiento adecuado y lograr el objetivo principal.

Figura 9. Proceso de manufactura de grandes rollos de papel Tissue



Fuente: (Proceso de Fabricación Del Papel, 2011)

Figura 10. Diagrama de proceso para la fabricación de grandes rollos de papel Tissue



Fuente: Completamente generada por el autor, 2023

2.1.1. PREPARACIÓN DE PASTA

Este proceso se encarga de convertir el papel reciclado en una pasta de celulosa limpia, libre de contaminantes.

Pulpeado: esta operación se la lleva a cabo en un equipo denominado pulper, el cual se encarga de desfibrar el material reciclado para liberar las fibras de celulosa, separar las partículas de tinta y otros contaminantes, mantenerlos en suspensión para su posterior eliminación.

Limpieza gruesa: operación en la cual se eliminan contaminantes de tamaño grande y mediano como plásticos, madera, trapos, grapas, piedras, etc. Para esto se utiliza limpiadores centrífugos de alta densidad y tamices de agujeros entre 1,6 y 6 mm de diámetro.

Des-tintado: proceso en el cual se elimina la tinta y otros contaminantes de tamaños 10 a 250 μ , se realiza en una celda de destintado.

Limpieza fina: se realiza con limpiadores centrífugos de baja densidad y tamices de ranuras entre 0,12 a 0,15 mm, en donde se elimina stickies, carbonatos, arena y puntos de tinta que no alcanzaron a flotar en la celda de destintado.

Espesado: todas las operaciones de limpieza de la fibra de celulosa se realizan a bajas concentraciones y con altos volúmenes de agua, una vez limpia la fibra es necesario eliminar cierta cantidad de agua para subir la concentración de fibras, esto se consigue con un espesador cilíndrico de malla mesh 100.

Dispersión en caliente: para mejorar la calidad de la pasta de celulosa que proviene del papel reciclado, especialmente el contenido de puntos de tinta, se utiliza el proceso de dispersión en caliente, por medio de temperatura de 70 a 90 °C, alta concentración de fibras, 25 a 30%, por medio de un kneader o disperger se consigue reducir los puntos a tamaños que ya no son percibidos por el ojo humano.

Blanqueo oxidativo: proceso químico que, a base de peróxido de hidrógeno, hidróxido de sodio y temperatura de 70 a 90 °C, ayuda a blanquear las fibras de celulosa (Caluña & Patricio, 2018).

2.1.2. MANUFACTURA DE GRANDES ROLLOS

La manufactura de grandes rollos se produce en la secadora de papel con la secuencia de las operaciones que se realiza durante la producción del papel; en la Figura 11 se puede ver el diagrama del proceso.

Mezcla de fibras: consiste en mezclar fibras de celulosa reciclada provenientes de la planta de preparación de pasta con fibras de celulosa virgen que viene directamente desde el proceso de desfibrado de forma automática según la formulación empleada.

Refinación: es la acción de las fuerzas mecánicas e hidráulica que se lleva a cabo sobre las fibras de celulosa, mediante un equipo denominado refinador, para alcanzar el desarrollo óptimo de las resistencias de las fibras, mantener el control de la formación de la hoja de papel y drenado del agua en la tela formadora.

Formación: es la transformación de las fibras de celulosa en suspensión, a una hoja fina de fibras entrecruzadas por medio de un inyector, una malla y un rodillo formador que gira a gran velocidad.

Prensado: la hoja formada tiene un alto contenido de agua, la cual es necesario extraer, para ello se utiliza un fieltro, el cual traslada la hoja desde la zona de formación hacia la zona de secado, pasando por una zona de prensado y cajas de vacío que eliminan la mayor cantidad de agua.

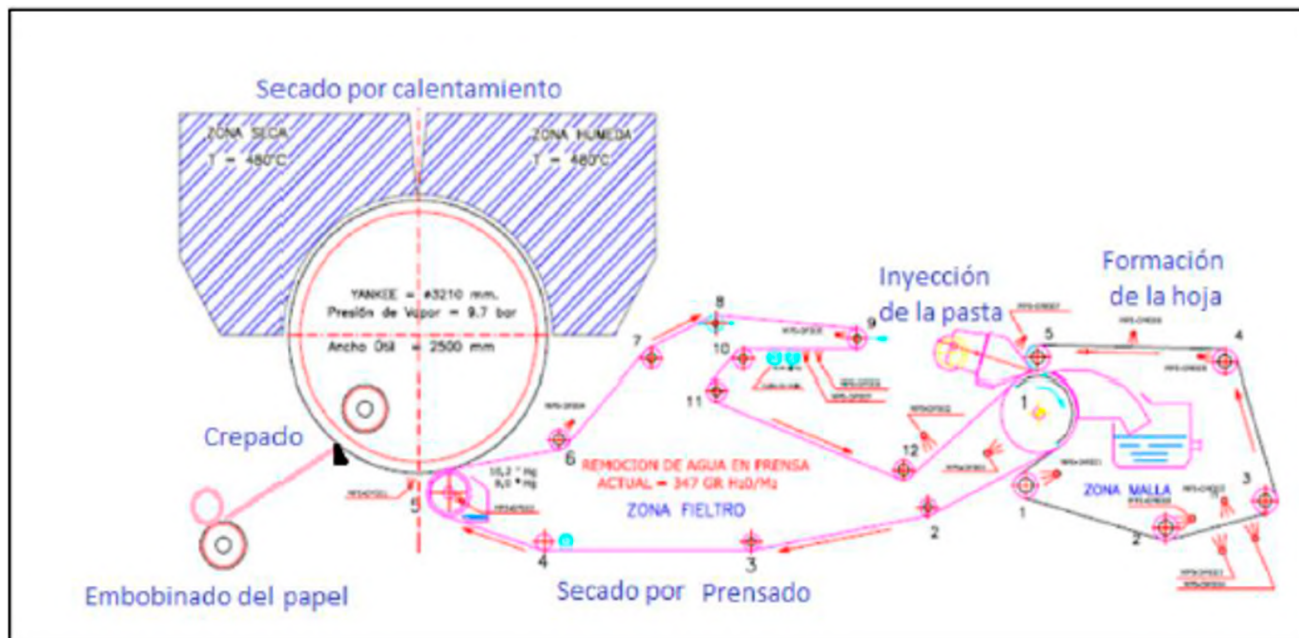
Secado: la hoja de papel que sale de la zona de prensado contiene un 40% de sequedad, la cual tiene que llegar al 95% de sequedad, para ello se utiliza un cilindro (yankee) presurizado con vapor que alcanza los 120 °C en la superficie externa que está en contacto con la hoja de papel, mientras la cara externa del papel es secada con aire caliente proveniente de una campana que sopla aire caliente de 450 °C, proveniente del Hood y ventiladores.

Crepado: operación encargada de separar la hoja de papel del cilindro secador (yankee), se lo realiza con la ayuda de una cuchilla crepadora que mantiene ángulos específicos de ataque al secador, además se utiliza un polímero coating que crea una película fina en la superficie del secador, para que el contacto de la hoja no sea directo con el metal, y un agente release, que es un aceite que mantiene lubricada la cuchilla y facilita el despegue de la hoja. Esta operación es la que le da las características propias del papel Tissue.

Bobinado: es la operación final donde el papel Tissue producido es bobinado en grandes rollos a velocidades entre 800 y 2.000 m/min según la tecnología de la máquina de papel.

Conversión: esta es la etapa final donde se transforma los grandes rollos producidos en los molinos en rollos pequeños como se conocen en el mercado (Caluña & Patricio, 2018; Fleiter et al., 2012; Laurijssen et al., 2010; Sivill et al., 2005; Trokhan, 1987).

Figura 11. Diagrama de operaciones del proceso de secado de papel Tissue

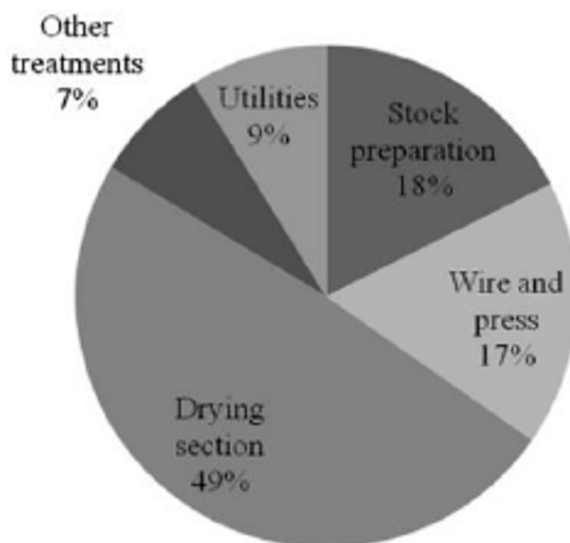


Fuente: (Caluña & Patricio, 2018)

2.1.3. CONSUMOS DE ENERGÍA EN EL PROCESO DE SECADO TISSUE

A nivel mundial la industria de papel es la cuarta más grande en consumo de energía eléctrica, son responsables del 6% del consumo de energía total a nivel mundial (Laurijssen et al., 2010). Las máquinas secadoras son una tecnología que varía muy poco dependiendo del fabricante, ya que esta tecnología continúa siendo prácticamente igual a su diseño original inventada por Wahren en 1970 (Valmet, 2013). El secado térmico a menudo es responsable del 50% del consumo total de energía y del 80% del consumo de vapor dentro de la planta de producción de papel Tissue y del 50% del consumo de energía eléctrica, en forma general el consumo de estas máquinas secadoras está alrededor del 48 a 52%, como se puede ver en la Figura 12, se observar que corresponde al 49% del consumo de energía. En la empresa analizada el proceso de secado corresponde al 50% del consumo de energía.

Figura 12. Distribución normal del uso de energía primaria en la industria de producción de papel



Fuente: (Laurijssen et al., 2010)

2.1.4. DATA MINING

La minería de datos es un concepto que empieza alrededor de los años sesenta, donde los estadísticos comienzan a explorar información almacenada sin una definición previa de base de datos. Es hasta principios de los años ochenta cuando se empiezan a consolidar los conceptos de Data mining y knowledge discovery in databases (KDD) por Rakesh Agrawal, Gio Wiederhold, Robert Blum y Gregory Piatetsky-Shapiro entre otros.

Esta tecnología ha sido una herramienta importante en la historia para personas que se desenvuelven en los ámbitos académicos y de negocios. Esta evolución se ha plateado en 4 etapas a través de la historia.

- Colección de Datos (1960).
- Acceso de Datos (1980).
- Almacén de Datos y Apoyo a las Decisiones (principios de la década de 1990).
- Minería de Datos Inteligente (finales de la década de 1990).

El dato es un hecho que describe un suceso, lo importante de los datos es su capacidad de asociarse en un contexto donde se convierte en información útil. Por sí mismos los datos no brindan un significado y mucho menos los comunican y por esta razón no alteran ningún comportamiento. Pero si esta información se interpreta de forma adecuada, permitirá tomar mejores decisiones en las organizaciones según su contexto informativo.

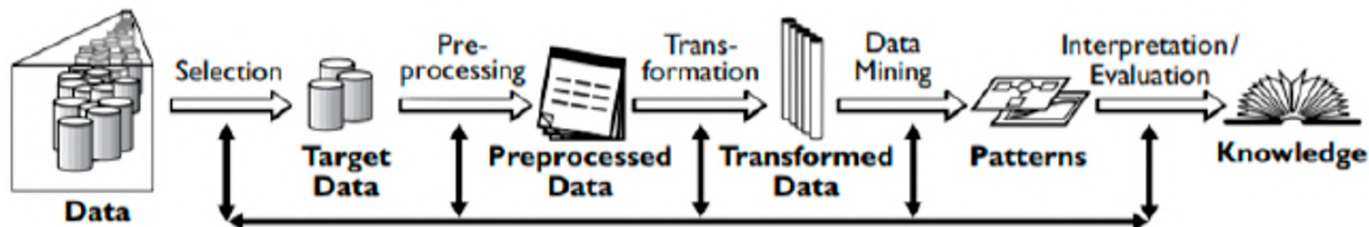
La minería de datos es el proceso de descubrir automáticamente información útil en grandes repositorios de datos. Las técnicas de minería de datos se implementan para explorar grandes bases de datos con el fin de encontrar patrones novedosos y útiles que de otro modo podrían seguir siendo desconocidos (Larose & Larose, 2014).

El proceso KDD se define como “el proceso no trivial de identificar patrones válidos, novedosos y potencialmente útiles y en última instancia, comprensibles a partir de los datos” según U. Fayyad, G. Piatetsky-Shapiro, y P. Smyth (Fayyad et al., 1996). El objetivo principal del KDD, es encontrar conocimiento útil, válido, relevante y nuevo sobre una determinada actividad mediante algoritmos, dadas las crecientes órdenes de magnitud en los datos.

Cabe destacar que el KDD tiene cierta convergencia con diferentes disciplinas de investigación tales como aprendizaje automático o machine learning (ML), estadística, inteligencia artificial, sistemas de gestión de base de datos, técnicas de visualización de datos, sistemas para el apoyo a la toma de decisión o la recuperación de información, entre otras.

El análisis de los datos establecido por KDD tiene las siguientes etapas (ver Figura 13).

Figura 13. Resumen de los pasos que componen el proceso KDD



Fuente: (Fayyad et al., 1996)

Selección de datos: consiste en buscar el objetivo y las herramientas del proceso de minería, identificando los datos que van a ser extraídos; esto quiere decir, primero se debe tener en cuenta lo que se quiere obtener y cuáles son los datos que nos facilitarán esa información para poder llegar a el objetivo de este trabajo.

Pre-procesamiento de los datos: en este paso se limpian los datos sucios, incluyendo los datos incompletos (donde hay atributos o valores de atributos perdidos), el ruido (valores incorrectos o inesperados) y datos inconsistentes (conteniendo valores y atributos con nombres diferentes). Los datos sucios en algunos casos deben ser eliminados ya que pueden contribuir a un análisis inexacto y resultados incorrectos.

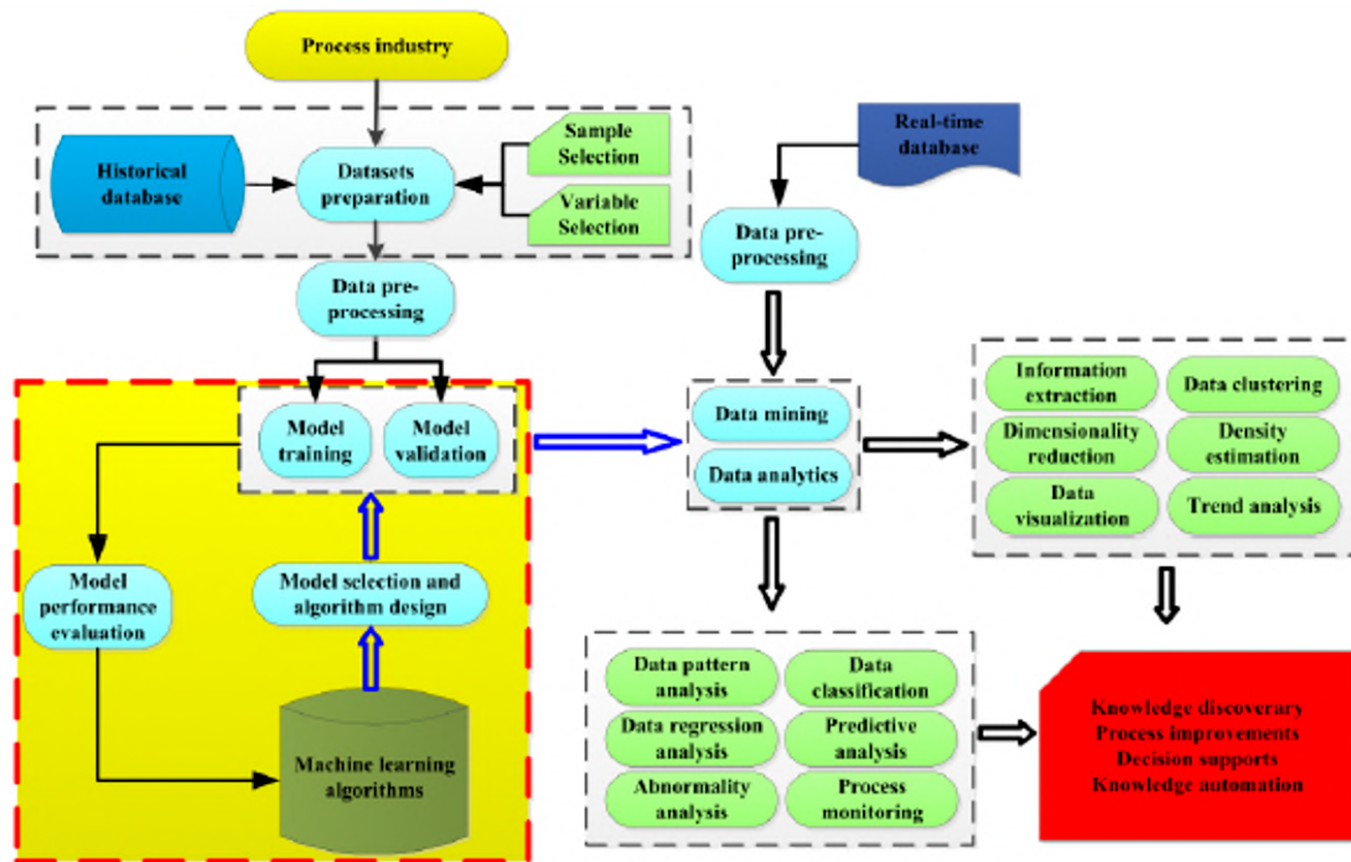
Transformación de datos: la transformación discreta de los datos tiene la ventaja de que mejoran la comprensión de las reglas descubiertas para transformar los datos de bajo nivel en datos de alto nivel. Lo anterior también permite reducir significativamente el tiempo de ejecución del algoritmo de búsqueda. Sin embargo, presenta una desventaja asociada a la posibilidad de reducir la exactitud del conocimiento descubierto, debido a que puede causar la pérdida de alguna información. Existen diferentes métodos de transformación de variables continuas a discretas que se pueden agrupar según distintas aproximaciones. Métodos locales (realizan la transformación discreta en una región del espacio de las instancias, por ejemplo, utilizando un subconjunto de las instancias), métodos globales (utilizan el espacio de las instancias), métodos supervisados utilizan la información de la clave (valor del atributo objetivo). Es decir, en esta etapa se mejora la calidad de los datos con transformaciones que involucren ya sea reducción de dimensionalidad (disminuir la cantidad de variables del conjunto de datos) o bien transformaciones como por ejemplo convertir los valores que son números a categóricos (discretización).

Patrones: en la minería de datos hay patrones interesantes que representan conocimiento usando diferentes técnicas, incluyendo análisis estadísticos, lenguajes de consultas e interpretación por expertos.

Interpretación y evaluación de resultados: consiste en entender los resultados del análisis y sus implicaciones, para de manera objetiva realizar la interpretación acertada del análisis.

Ge. Zhiqiang, Song Zhihuan, Ding Steven X, Huang Biao (Ge et al., 2017), proponen la siguiente metodología para la minería de datos (ver).

Figura 14. Visión general de la metodología de análisis y minería de datos



Fuente: (Ge et al., 2017)

En resumen, DM es la etapa del descubrimiento en el KDD, usa algoritmos concretos que generan una enumeración de patrones a partir de los datos preprocesados (Han et al., 2011). Mientras la minería de datos tiene una función exploratoria para descubrir patrones, el ML se usa para reproducir patrones conocidos y hacer predicciones basadas en los patrones.

2.1.5. MACHINE LEARNING

El aprendizaje automático o Machine Learning (ML) es un subcampo de la informática que se desarrolló a partir del estudio del reconocimiento de patrones y la teoría del aprendizaje computacional en la inteligencia artificial. Sin embargo, la teoría de la probabilidad y las estadísticas también juegan un papel muy importante en el aprendizaje automático moderno. El aprendizaje automático explora el estudio y la construcción de algoritmos que pueden aprender y hacer predicciones sobre los datos. En 1959, Arthur Samuel definió el aprendizaje automático como un “campo de estudio que otorga a las computadoras la capacidad de aprender sin ser programado explícitamente” (Ge et al., 2017; Simon, 2013), y más tarde Tom M. Mitchell proporcionó una definición más formal y ampliamente citada: “A se dice que el programa de computadora aprende de la experiencia E con respecto a alguna clase de tareas T y la medida de rendimiento P, si su desempeño en las tareas en T, medido por P, mejora con la experiencia E” (Michalski et al., 2013).

En la actualidad la tendencia industrial viene dada por una revolución de fábrica inteligente, un concepto que nace de industria 4.0 o cuarta revolución industrial, esta consiste en la introducción de tecnologías digitales con la que se almacenan datos que posteriormente pueden ser procesados para analizarse.

El ML tiene dos principales métodos para aplicaciones industriales, el aprendizaje supervisado y el aprendizaje no supervisado, estos comprenden el 80-90% de las aplicaciones. De los dos métodos anteriormente mencionados, se derivan las estrategias semi-supervisadas, rara vez usadas en el área de procesamiento de datos a industrias.

2.1.6. MÉTODO DE APRENDIZAJE NO SUPERVISADO

El aprendizaje no supervisado se utiliza en datos que no tienen etiquetas. El objetivo principal del método de aprendizaje no supervisado es explorar los datos y encontrar alguna estructura oculta entre ellos. Para las aplicaciones en la industria, este

tipo de métodos de aprendizaje automático se utilizan principalmente en reducción de la dimensionalidad, extracción de información, visualización de datos, estimación de la densidad, detección de valores atípicos, monitoreo de procesos, entre otros. (Ge et al., 2017; Hastie et al., 2009).

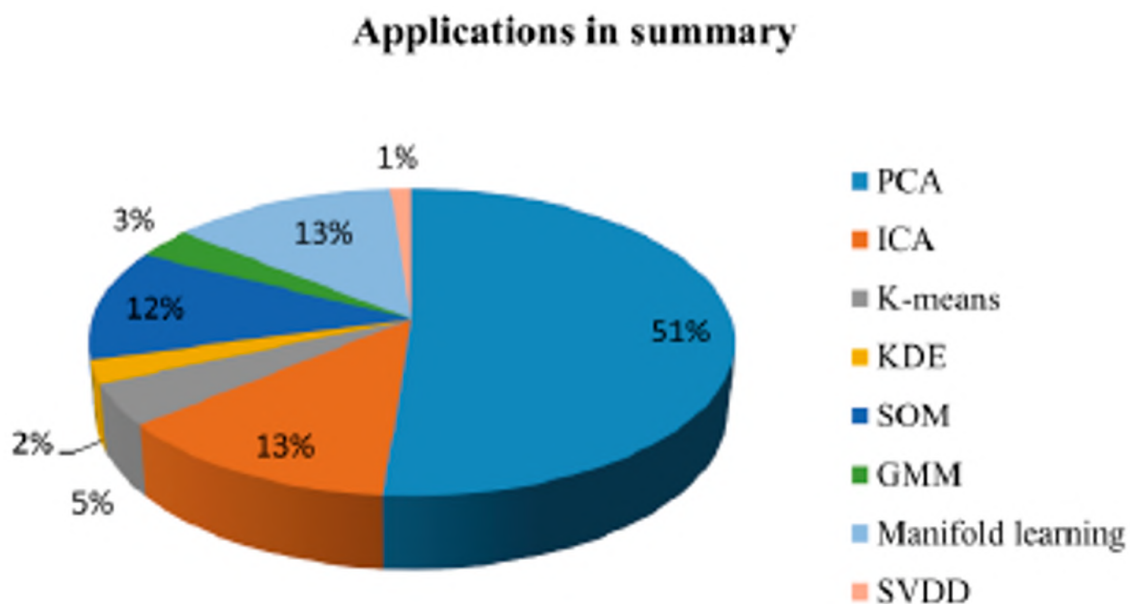
Los métodos de aprendizaje aplicados normalmente en la industria son:

1. Análisis de componentes principales en inglés Principal Component Analysis (PCA).
2. Análisis de componentes independientes en inglés Independent Component Analysis (IAC).
3. K-método de agrupamiento en inglés K-means clustering.
4. Estimación de la densidad del Kernel en inglés Kernel density estimation (KDE).
5. Mapa autoorganizado en inglés Self-organizing map (SOM).
6. Modelo de mezcla Gaussiana en inglés Gaussian mixture model.
7. Método de aprendizaje variado o múltiple en inglés Manifold Learning Method.
8. Descripción de los datos del vector de soporte en inglés support vector data description.

En resumen, los métodos de aprendizaje no supervisados se han utilizado ampliamente en la industria de procesos en las últimas dos décadas. La mayoría de las aplicaciones del método no supervisado se realizaron a través del modelo PCA. El método anteriormente mencionado proporciona un estado general de la aplicación sobre los 8 métodos principales de aprendizaje sin supervisión entre 2000 y 2015. Se puede ver en la Figura 15 que las aplicaciones basadas en el modelo de PCA representan más de la mitad de todas las aplicaciones. En los últimos años, la introducción acelerada de múltiples algoritmos de aprendizaje ha incrementado las aplicaciones de los métodos no supervisados en la industria de procesos, representando un 12% de todas las aplicaciones a nivel industrial.

En particular se tiene aplicaciones detalladas de todos los métodos de aprendizaje no supervisados en términos de reducción de la dimensión, detección de valores atípicos, monitoreo de procesos y visualización de datos. Se puede ver que el PCA juega un papel dominante en cada aspecto de la aplicación. Para la reducción de la dimensionalidad, el método de aprendizaje múltiple también ha servido para muchas aplicaciones, debido a su gran capacidad en la reducción de la dimensionalidad no lineal. El método de mapa auto-organizado ha proporcionado un número de solicitud comparativo en términos de visualización de datos, en comparación con el PCA. Esto se debe principalmente a que el mapa auto-organizado describe un mapeo desde un espacio de datos de alta dimensión a un espacio de datos de baja dimensión, lo que hace que la visualización de datos sea más conveniente para el proceso (Aldrich & Auret, 2013; Ge et al., 2017; Jolliffe, 2011; Michalski et al., 2013; Shang et al., 2017; Zhao et al., 2018; Zhu et al., 2017).

Figura 15. Métodos más aplicados en el aprendizaje no supervisado



Fuente: (Ge et al., 2017)

2.1.7. MÉTODO DE APRENDIZAJE SUPERVISADO

A diferencia del método de aprendizaje no supervisado, el aprendizaje supervisado se ocupa de muestras de datos etiquetadas, discretas o continuas. Cuando la etiqueta de la muestra de datos es un valor discreto, el aprendizaje supervisado se puede usar para clasificar los datos del proceso, por ejemplo: clasificación de fallas o clasificación de modo de funcionamiento. De lo contrario, cuando la etiqueta de la muestra de datos es un valor continuo, los modelos de regresión se pueden construir con fines de predicción y estimación.

Las principales aplicaciones del método de aprendizaje automático supervisado incluyen monitoreo de procesos, clasificación e identificación de fallas, localización de modo de operación en línea, modelado de sensores suaves y aplicaciones en línea, predicción de calidad y estimación en línea, predicción y diagnóstico de índices de rendimiento clave, entre otros (Ge et al., 2017; Quinlan, 2014).

1. Regresión de componente principal en inglés Principal component regression (PCR).
2. Regresión de mínimos cuadrados parciales en inglés Partial least squares regression (PLS regression).
3. Análisis discriminante de pescadores en inglés Fisher Discriminant Analysis (FDA).
4. Regresión lineal multivariable en inglés Multivariate linear regression.
5. Redes neuronales artificiales en inglés Artificial neural networks (ANNs).
6. Máquinas de vectores soporte en inglés Support vector machine (SVM).
7. Algoritmo de vecinos más cercanos en inglés Nearest neighbors algorithm.
8. Regresión del proceso Gaussiano en inglés Gaussian Process Regression.
9. Árbol de decisión en inglés Decision Tree.
10. Bosque al azar en inglés Random Forest.

Para resumir los métodos de aprendizaje automático supervisado utilizados en la industria de procesos desde el año 2000.

El aprendizaje supervisado ofrece una descripción general del estado de la aplicación para 10 métodos de aprendizaje supervisado diferentes, los resultados de estos métodos más relevantes son las redes de neuronales artificiales y los mínimos cuadrados parciales estos dos métodos son los más utilizados en el mundo.

En los últimos 15 años, estos dos métodos representan más de la mitad de las aplicaciones en la industria de procesos. Más específicamente, la aplicación de todos los métodos de aprendizaje supervisados en términos de monitoreo de procesos, clasificación de fallas, sensor de software y predicción de calidad. Con el propósito de monitorear el proceso, PLS y ANN han jugado roles dominantes, debido

a la gran capacidad de SVM en la clasificación de datos, ha obtenido la mayoría de las aplicaciones en clasificación de fallas. Para desarrollos de sensores suaves y predicciones de calidad, PLS y ANN son también los dos métodos más populares que ocupan aproximadamente el 60% de todas las aplicaciones Figura 16 (Bottou et al., 2018; Câmara et al., 2017; Ge et al., 2017; Han et al., 2011; Jolliffe, 2011; Nemeth & Michalconok, 2017; Quinlan, 2014).

Los ejemplos más claros de análisis supervisados son los que se utilizan para crear y desarrollar algunas aplicaciones empresariales de uso común como las siguientes:

Reconocimiento de imágenes: estos algoritmos son ampliamente usados en él para aislar y reconocer objetos de imágenes o videos lo cual resulta útil en la visión computacional para el análisis de imágenes.

Análisis de opinión del cliente: mediante el uso de algoritmos de machine learning supervisado, las organizaciones pueden extraer y clasificar información importante de grandes volúmenes de datos, incluidos contexto, emoción e intención, con muy poca intervención humana. Esto puede ser increíblemente útil para comprender mejor las interacciones del cliente y puede usarse para mejorar las iniciativas de compromiso con la marca.

Detección de correo no deseado: la detección de correo no deseado es otro ejemplo de modelo de aprendizaje supervisado. Mediante algoritmos de clasificación supervisados, las organizaciones pueden entrenar bases de datos para reconocer patrones o anomalías en nuevos datos para organizar correspondencias relacionadas con correo no deseado de manera efectiva.

Análisis predictivo: un caso de uso generalizado de los modelos de aprendizaje supervisado es la creación de sistemas de análisis predictivo para proporcionar información más exhaustiva sobre varios puntos de datos de negocio. Esto permite a las empresas anticipar determinados resultados con base en una determinada variable de salida, ayudando a los líderes de la empresa a justificar decisiones o realizar cambios en beneficio de la organización (IMB, 2013). Este proyecto es de análisis predictivo, debido a que los líderes de la empresa de fabricación de papel establecen cuáles son las necesidades puntuales para mejorar el proceso productivo y tomar decisiones basadas en análisis de datos supervisados.

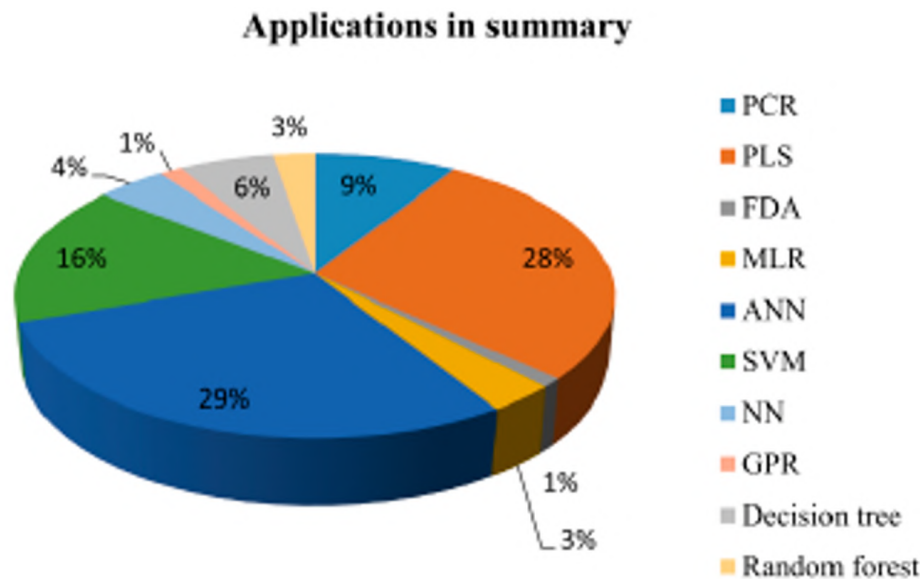
Con base en las necesidades de conocer una predicción del costo con un buen rendimiento y adicionalmente poder conocer los valores que hacen mínimo el costo productivo es necesario implementar el uso de los algoritmos de aprendizaje supervisado, con la investigación previa se puede determinar que los algoritmos inteligentes más usados y con mejores resultados en la industria en problemas de análisis predictivo de regresión son las Redes Neuronales sin embargo para lograr el objetivo de este trabajo es necesario conocer las variables que hacen mínimo el costo y estos valores se pueden conocer con un algoritmo como lo es la Regresión Lineal Múltiple.

Redes neuronales: se utilizan principalmente para los algoritmos de deep learning, procesan los datos de entrenamiento imitando la interconectividad del cerebro humano a través de capas de nodos. Cada nodo está formado por entradas, ponderaciones, un sesgo (o umbral) y

una salida. Si ese valor de salida excede un umbral determinado, “dispara” o activa el nodo, pasando datos a la siguiente capa de la red. Las redes neuronales aprenden esta función de correlación a través del aprendizaje supervisado y se ajustan con base en la función de pérdida a través del proceso de su gradiente descendente (Menacho Chiok, 2014; Moolayil et al., 2019).

Regresión lineal Múltiple: se utiliza para identificar la relación entre una variable dependiente y una o más variables independientes y, por lo general, se aprovecha para realizar predicciones sobre resultados futuros. Cuando solo hay una variable independiente y una variable dependiente, se conoce como regresión lineal simple. A medida que aumenta el número de variables independientes, se habla de regresión lineal múltiple. Para cada tipo de regresión lineal, trata de trazar la línea que mejor se ajuste, la cual se calcula mediante el método de mínimos cuadrados. Sin embargo, a diferencia de otros modelos de regresión, esta línea es recta cuando se traza en un grafo (Rojo Abuín, n.d.).

Figura 16. Métodos más aplicados en el aprendizaje supervisado



Fuente: (Ge et al., 2017)

2.2. RECOLECCIÓN DE DATOS

En esta segunda fase se hace una recolección de datos inicial, lo más completa posible del proceso específico de secado de papel. En esta fase se buscan los primeros contactos con los datos para familiarizarse con ellos, verificar su calidad, identificar relaciones evidentes para formular las hipótesis iniciales.

Los registros de los servidores están disponibles en formato CSV (valores separados por comas) de Excel. Para todo el procesamiento de datos requerido y entrenamiento de algoritmos se planeó trabajar con un lenguaje multiplataforma de código abierto (gratuito). Se seleccionó Python y se definió el potente entorno de desarrollo Spyder (Scientific Python Development Environment). Con base en Python se han desarrollado funciones avanzadas a través de librerías especializadas para el análisis de datos que fueron utilizadas en el proyecto, como lo son:

- Para Visualización: Matplotlib, Seaborn.
- Para Cálculo Numérico y Análisis de Datos: NumPy, SciPy, Pandas.
- Para Machine Learning: scikit-learn.
- Para Deep Learning: TensorFlow, Keras.

En la elección de las variables de interés, después de numerosas visitas al proceso industrial y análisis con el personal (expertos del proceso), se encontró que faltaban variables consideradas de interés de las cuales no se llevaba registro en el servidor, esta situación conllevó a esperar 3 meses para poder recolectar los datos nuevos, este periodo de tiempo está determinado por la capacidad de los servidores en su almacenamiento y el gran volumen de datos e información a respaldar en la empresa. De igual forma también se pudo detectar la duplicidad en el registro de algunas variables las cuales irán siendo objeto de un trabajo de depuración. Finalmente se tuvo en cuenta las variables que el operario puede manipular en el proceso de secado de forma directa (manualmente) debido a su impacto de gran relevancia en el proceso.

Los datos a registrar se ubican en cuatro secciones: dos secciones asociadas al proceso de mayor consumo, que es el de gas (sección Quemador lado Húmero con un total de 15 variables; sección Quemador lado seco, total 14 variables), otra al consumo de vapor (total 13 variables) y la última asociada al consumo de energía eléctrica (total 21 variables). Las variables en

forma general son Velocidades, Temperaturas, Fechas, Presiones, Potencias eléctricas, Flujos, Categóricas con el Nombre del Producto de fabricación, Pesos base (g/m^2) y los Consumos de energía eléctrica, gas y vapor. Para el tiempo de 3 meses, se almacenan los datos de forma matricial, donde se obtienen 63 columnas y 17015 filas, lo cual corresponde a un muestreo cada 5 minutos; la frecuencia se determinó debido a que no hay muchas variaciones en ese rango de tiempo por la estabilidad del proceso.

2.3. PREPROCESAMIENTO DE DATOS Y GENERAR DATA SET

En el preprocesamiento específico para los datos registrados del proceso fueron integrados algoritmos de normalización para las variables, una nueva selección y depuración de variables relevantes por parte de expertos y la generación de una nueva variable a partir de las ya registradas para cuantificar el costo de producción. Posteriormente, se realiza una separación de datos para eliminar los asociados a productos que no son el caso de interés y además, eliminar los datos de cambio de producto para que no afecten el análisis y solo quedarse con los datos de la operación real. Para la normalización de cada variable, con base en el histórico registrado, la Ecuación 1 expresa la fórmula utilizada, donde X_{new} es el valor normalizado de la variable, X es el valor de la variable, y X_{min} y X_{max} son los valores mínimo y máximo de la variable en el histórico de datos.

Ecuación 1. Función de normalización.

$$X_{new} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

En cuanto a selección y depuración de variables tenemos que: a) de las dos secciones de gas, de la sección Quemador lado Húmedo con un total de 15 variables fueron seleccionadas 2, y de la sección Quemador lado seco de un total 14 variables fue seleccionada 1; b) de la sección de consumo de vapor de 13 variables fueron seleccionadas 5; y c) de la sección de consumo de energía eléctrica de un total 21 variables fueron seleccionadas 7. En detalle, se relacionan las variables en la Tabla 2.

Tabla 2. Variables Seleccionadas

#	SECCIÓN	EDUCACIONAL	INSTRUCCIONAL
1	GAS (Quemador lado húmedo)	Gas_m3/h_LH'	Consumo de gas del quemador del Lado húmedo
2	GAS (Quemador lado seco)	Gas_m3/h_LS'	Consumo de gas quemador Lado Seco
3	Temperatura (Quemador lado húmedo)	T_Set_Point_Q_TT403B LH	Temperatura de set point Quemador Lado Húmedo, variable de manipulación manual que el operario puede manipular
4	VAPOR	P_Yank_SP_PT094	Presión del yankee Set Point; variable de manipulación manual que el operario puede manipular
5	VAPOR	Ape_TermComp_P_DIF_PDV096A	Apertura termocompresor presión diferencial, variable de manipulación manual que el operario puede manipular
6	VAPOR	P DIF Yankee PIDC096_PV	Presión diferencial del Yankee, variable de manipulación manual que el operario puede manipular
7	VAPOR	P_DIF_SP_YANK_PDIC_SP	Presión diferencial yankee, variable de manipulación manual que el operario puede manipular
8	VAPOR	kw/ton E V SaI_Yank_FT096	kw Tonelada energía vapor salida yankee, consumo de vapor
9	ELECTRICA	KW_VEN001	Consumo de electricidad ventilador, variable de manipulación manual que el operario puede manipular
10	ELECTRICA	KW_VEN002	Consumo de electricidad ventilador, variable de manipulación manual que el operario puede manipular
11	ELECTRICA	KW_VEN003	Consumo de electricidad ventilador, variable de manipulación manual que el operario puede manipular
12	ELECTRICA	KW_VEN004	Consumo de electricidad ventilador, variable de manipulación manual que el operario puede manipular
13	ELECTRICA	KW_VEN005	Consumo de electricidad ventilador, variable de manipulación manual que el operario puede manipular
14	ELECTRICA	KW_VEN012	Consumo de electricidad ventilador, variable de manipulación manual que el operario puede manipular
15	ELECTRICA	M/min_YK001A	Consumo de electricidad del Yankee, variable de manipulación manual que el operario puede manipular

Fuente: Completamente generada por el autor, 2023

En cuanto a la depuración y selección de variables, son fundamentales las variables que puede manipular el operario por cuanto afectan el costo del producto, es decir estas manipulaciones afectan los consumos de gas, vapor, y electricidad; además, que como referencia el estudio se direcciona a optimizar el proceso de mayor fabricación dentro de la empresa, con el objetivo de encontrar variables óptimas de trabajo para el mínimo costo y un predictor del costo con alto rendimiento.

En cuanto al cálculo de una nueva variable, para calcular el costo (en COP) del consumo de energía, de acuerdo con la Ecuación 2, se realiza la sumatoria de tres términos asociados a los consumos de gas, electricidad y vapor. El primer término tiene en cuenta el costo en COP del gas multiplicado por consumo resultante de GasLH y GasLS; el segundo término requiere de la suma de las 7 variables de consumo de electricidad multiplicado por el costo en COP de la electricidad; finalmente, en el tercer término, la variable de consumo de vapor ya está unificada y solo resta multiplicarla por el costo de esta energía correspondiente. El costo total, es adicionado a data set, quedando finalmente una matriz de datos de 6321 registros con 13 variables.

Ecuación 2. Cálculo de costo total de consumo de gas, vapor y electricidad.

$$Costo_{Total} = Y = [Gas_{LH} + Gas_{LS}] * Costo_{gas} + [VEN_{001} + VEN_{002} + VEN_{003} + VEN_{004} + VEN_{005} + VEN_{012} + YK_{001}] * Costo_{electricidad} + [Vapor Yank_{FT096}] * Costo_{vapor}$$

Para el caso en estudio los preprocesamientos son aplicados buscando solo analizar el producto de mayor fabricación en la empresa, llegando a obtener datos completamente limpios sin valores faltantes ni nulos, es decir 0% de outliers.

2.4. ENTRENAMIENTO DE ALGORITMO

Los algoritmos de Machine Learning, son sistemas que consisten en el ajuste de parámetros según los datos de entrada para la detección automática de patrones relevantes de un conjunto de datos (Maisueche Cuadrado, 2019). Los tipos de aprendizaje de Machine Learning se categorizan principalmente en aprendizaje supervisado y aprendizaje no supervisado, ambos son ampliamente usados en la industria; este proyecto es de aprendizaje supervisado, donde se conocen directamente las variables o descriptores del proceso tomadas para un solo producto de fabricación (Pedregosa et al., 2011). Para la evaluación de los modelos se utilizó el criterio R-Square (R^2), ver Ecuación 3.

Ecuación 3. R al cuadrado (R^2)

$$R^2 = 1 - \frac{MSE(modelo)}{MSE(baseline)} = 1 - \frac{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$$

R^2 es una medida estadística que establece lo cerca que están los datos de la línea de regresión ajustada. También se conoce como coeficiente de determinación o coeficiente de determinación múltiple si se trata de regresión múltiple, está dado en porcentaje y entre más cercano a 100% mejor se ajusta el modelo a los datos, esta medida estadística es usada por la librería scikit learn de Python.

2.4.1. REGRESIÓN LINEAL MÚLTIPLE

La regresión lineal múltiple es un método estadístico que trata de ajustar modelos lineales o linealizables entre una variable dependiente y más de unas variables independientes, en Python el mismo modelo de regresión lineal simple se encarga de realizar las tareas del modelo de regresión múltiple (Ge et al., 2017; Menacho Chiok, 2014). Para el desarrollo de este proyecto este método proporciona dos análisis, cuanto es el Score R^2 , de un método estadístico y la optimización de variables para conocer el valor mínimo operativo donde se obtendrá el mínimo costo.

Ecuación 4. Modelo de Regresión Lineal Múltiple.

$$Y = a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + a_6x_6 + a_7x_7 + a_8x_8 + a_9x_9 + a_{10}x_{10} + a_{11}x_{11} + a_{12}x_{12} + b$$

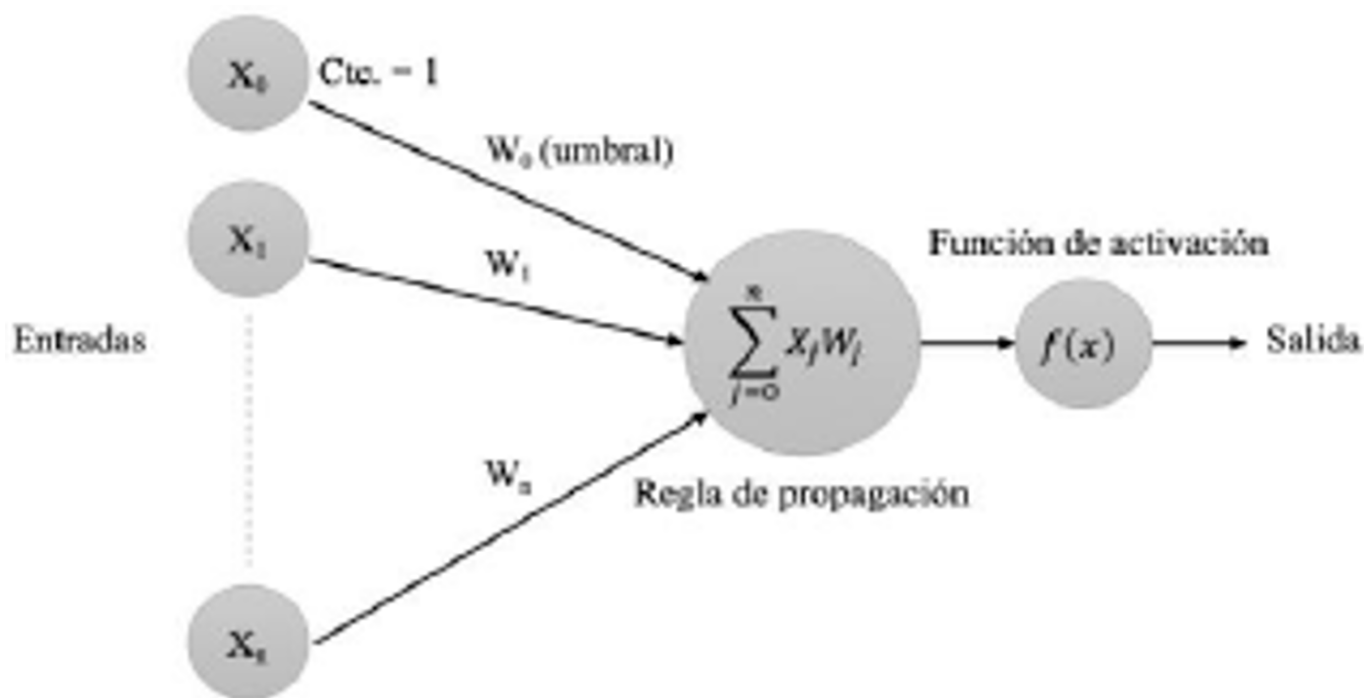
Donde **Y** es la variable dependiente (**CostoTotal**, **an** corresponde a los coeficientes o valores de pendiente (ver Tabla 3) donde está el resultado de los coeficientes), **xn** son los valores de las variables independientes y **b** es el intercepto.

2.4.2. REDES NEURONALES

Las Redes Neuronales Artificiales son una herramienta de modelización de datos potente, con el principio de funcionamiento de la neurona humana, esto le permite desarrollar tareas complejas a través del aprendizaje adaptativo. Sus aplicaciones se dan en muchos campos como en la medicina, las finanzas, procesamiento de señales, diagnósticos, control de robótica, visión artificial y muchos otros más. Los algoritmos convencionales no compiten con las redes neuronales, se complementan con estas, dependiendo de las tareas es más apropiado un enfoque algorítmico o un enfoque neuronal, las Redes Neuronales son de gran utilidad en problemas con variables categóricas y visión artificial (Maarouf, 2016). El potencial más grande de las redes Neuronales es aprender y aproximar relaciones entre las variables de entrada y salida, independientemente de la complejidad y el tamaño del problema (Maldonado & Acevedo, n.d.; Menacho Chiok, 2014; Tabares-Ospina & Hernández-Riveros, 2009).

Se elige trabajar con el modelo de neurona artificial basado en redes neuronales, correspondiente a una red neuronal artificial supervisada del tipo Perceptrón Multicapa, con la alimentación hacia delante de las variables anteriormente mostradas y con entrenamiento/aprendizaje para la propagación hacia atrás, se realizó como variable de salida el costo del producto, ya que es el modelo más utilizado y con mejores resultados para optimizar procesos industriales (Ge et al., 2017). Se eligen las funciones de Python ya que son modelos altamente probados y optimizados que garantizan gran fiabilidad. A continuación, se ilustra una Red Neuronal Artificial estándar.

Figura 17. Red neuronal artificial estándar



Fuente: (Meireles et al., 2003)

En forma general no existe una sistematización de cómo se implementan estas redes, en cuanto a entradas, número de capas ocultas en la red multicapa y número de datos para entrenamiento y validación, debido a que esto es dependiendo de las características del sistema en particular. Los modelos de Redes Neuronales que cubren el 90% de las aplicaciones son MLP, SOF y RBF, encontrando que las MLP tienen el 81.2% (Edwards et al., 1999; Haque & Kashtiban, 2000; Meireles et al., 2003).

La neurona Estándar está definida por lo siguiente:

Entrada: son las variables seleccionadas en las que influye el operario manualmente (velocidades, temperaturas, presión, potencia eléctrica, flujo y peso base).

Pesos sinápticos: representan la intensidad de interacción entre cada neurona.

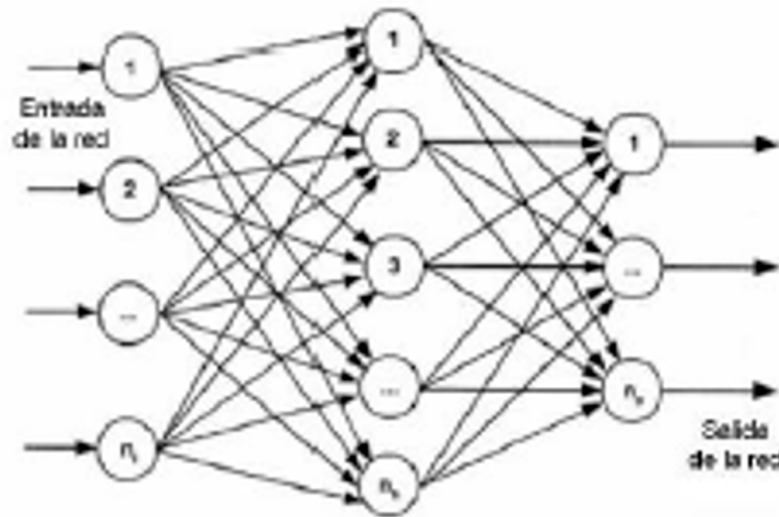
Umbral: entrada adicional que la neurona debe sobrepasar para activarse.

Función de Activación: es un algoritmo o función matemática que procesa las entradas previamente multiplicadas por los pesos y sumadas con el umbral. Las principales funciones de activación son Escalón, Lineal, Lineal a Tramos, Sigmoidal y Gausiana.

Salida: Es la variable objetivo en el caso de este proyecto es el costo final del producto.

El perceptrón multicapa o red multicapa con conexión hacia adelante en la actualidad es una de las arquitecturas más utilizadas en la resolución de problemas. Esto es debido, fundamentalmente, a su capacidad como aproximador universal, así como a su fácil uso y aplicabilidad.

Figura 18. Red neuronal. MPL (Multi Layer Perceptron)

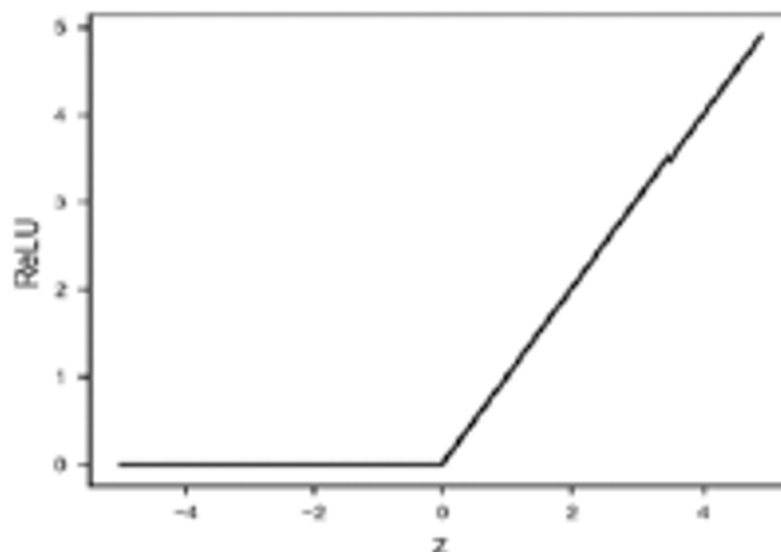


Fuente: (Meireles et al., 2003)

Se realizará una tabla con 20 iteraciones en donde los principales cambios iterativos se darán en el número de epochs, batch size y hidden layer, también se harán cambios de algunos parámetros como los que se describen más adelante. La literatura encontrada recomienda que para problemas de regresión no es conveniente algunos cambios descritos posteriormente ya que el modelo matemático no es capaz de predecir el comportamiento como lo es el caso de funciones logarítmicas o exponenciales.

Función de activación ReLU: la función puede parecer lineal, pero no lo es. ReLU es una función no lineal válida y de las más usadas como función de activación de hecho. Esta función no solo mejora el rendimiento, sino que ayuda significativamente a reducir el número de cálculos durante la fase de entrenamiento. En la siguiente Figura 19 se ilustra la función.

Figura 19. ReLU Función de activación



Fuente: (Moolayil et al., 2019)

Modelo: el modelo a usar es el 'Sequential' de Keras.

Capas: se realizarán una serie de iteraciones para definir el número de capas.

La Función de Pérdida (The Loss Function): es la métrica que ayuda a una red a comprender si está aprendiendo en la dirección correcta. Para los problemas que no son categóricos se recomienda (Mean Absolute Error o Mean Squared Error). En el caso de este proyecto es un problema de regresión en donde la predicción final sería un número Costo final del producto en producción. Para iniciar todos los modelos se evaluarán con la fusión de pérdida 'Mean Absolute Error' y una vez encontrado la mejor iteración se realizará la evaluación de Mean Squared Error.

Optimizadores: la función de optimización es un algoritmo matemático que usa derivadas, derivadas parciales y regla de la cadena en calculo para comprender cuánto cambio se verá en la red de la función de pérdida al hacer un pequeño calculo en la neurona. La función de pérdida ayuda a determinar.

Para este proyecto se elige 'Adam', este es el optimizador más usado y recomendado según la literatura investigada. Esta técnica de optimización calcula una tasa de aprendizaje adaptativo para cada parámetro. Define el impulso y la varianza del gradiente de la pérdida y aprovecha un efecto combinado para actualizar los parámetros de peso.

Ecuación 5. Representación matemática del optimizador Adam.

Weights=Weights-(Momentum and Variance Combined)

Métricas: las métricas pueden entenderse como la función utilizada para juzgar el rendimiento del modelo en un conjunto de datos. La única diferencia entre las métricas y la función de pérdida es que los resultados de las métricas no se utilizan para entrenar el modelo con respecto a la optimización. Solo se utilizan para validar los resultados de la prueba. Para este proyecto se utilizará 'metrics.binary_accuracy, mean_absolute_percentage_error, Mean Absolute Error'.

Modelo de entrenamiento: una vez configurado el modelo de entrenamiento, es necesario ir variando los valores de cada época (epoch) esta variable permite evaluar si el modelo se está desempeñando como se desea con los valores de validación. El modelo aprovecha los datos de entrenamiento para entrenarse y aprender los patrones y al final de cada época, usará los datos de validación no vistos para hacer predicciones y calcular métricas. El rendimiento en el conjunto de datos de validación es una buena señal para el rendimiento general. Otro parámetro importante para la evaluación del modelo es la iteración de capas de profundidad de la neurona y ver el comportamiento de aprendizaje hasta encontrar las capas optimas de aprendizaje para el modelo en particular.

Finalmente, el parámetro tamaño de lote (batch size), funciona eligiendo un tamaño de lote de 12. Por lo tanto, si hubiera 6000 muestras de entrenamiento, el modelo tomaría y procesaría 12 muestras a la vez en un lote antes de actualizar los pesos del modelo (Hurtado & Aguado, 2005; Maarouf, 2016; Moolayil et al., 2019; Tabares-Ospina & Hernández-Riveros, 2009).

3. RESULTADOS

Para lograr el objetivo de este trabajo se hace uso del método estadístico regresión lineal múltiple, donde se obtienen los valores de las variables optimas de operación, haciendo uso de la función Optimaze de Phytion se minimiza el costo encontrando los valores de las variables Optimas del proceso (Tabla 4).

El proceso del código consiste importar las librerías que se van a usar (Pandas, Numpy, Matplotlib, Sklearn, scipy), luego se definen las variables de análisis x (todas menos el costo) y se define la variable objetivo como Y (costo del producto). Se separa los datos de entrenamiento perteneciendo al 80% y con el 20% restante se comprueba el Score R^2 del modelo de entrenamiento y se define el modelo o algoritmo a utilizar (linear_model.LinearRegression). Se entrena el modelo y se calcula el Score buscando optimizar el resultado al máximo score posible finalmente el entrenamiento permitió obtener un resultado máximo $R^2=65\%$ (Tabla 3).

Luego del entrenamiento este algoritmo devuelve una ecuación con los coeficientes y el intercepto buscando predecir el costo Y, al entregarle los valores x, se definen los rangos de la operación (ver Tabla 2) y finalmente se calcula el vector resultado para obtener el mínimo costo (ver Tabla 3).

Tabla 3. Límites operativos de fabricación del papel

VARIABLE	RANGO
T_Set_Point_Q_TT403B LH	[410 - 450]
P_Yank_SP_PT094	[113 - 117]
Ape_TermComp_P_DIF_PDV096A	[60 - 80]
P DIF Yankee PIDC096_PV	[10 - 25]
P_DIF_SP_YANK_PDIC_SP	[10 - 20]
RPM_VEN001	[1530 - 1570]
RPM_VEN002	[1500 - 1530]
RPM_VEN003	[1270 - 1645]
RPM_VEN004	[2740 - 3000]
RPM_VEN005	[2670 - 2790]
RPM_VEN012	[915 - 925]
M/min_YK001A	[1520 - 1562]

Fuente: Completamente generada por el autor, 2023

Tabla 4. Tabla resultados algoritmo de regresión Lineal Múltiple

REGRESIÓN LINEAL MÚLTIPLE	
Intercepto Coef b	893773,29
Valor pendiente a1	1005,42
Valor pendiente a2	-1402,81
Valor pendiente a3	-465,24
Valor pendiente a4	170,92
Valor pendiente a5	-2633,48
Valor pendiente a6	-313,61
Valor pendiente a7	283,61
Valor pendiente a8	-4,91
Valor pendiente a9	-20,43
Valor pendiente a10	2,68
Valor pendiente a11	-167,88
Valor pendiente a12	91,89
R ²	65,0%

Fuente: Completamente generada por el autor, 2023

En el Tabla 4, se presenta la lista de variables con sus valores estadísticos promedio y los valores resultados obtenidos vía optimización (buscando el mínimo costo de producción), donde se puede apreciar que trabajar con los valores promedio es menos rentable ya que el COSTO PRODUCTO POR HORA (ver en la línea final del Tabla 4) es más costoso. Este resultado, permite validar con base en un valor de desempeño energético, que el conocimiento de los expertos en conjunción con la aplicación del método estadístico de Regresión lineal múltiple y el cálculo de optimización, permitió conocer las variables y rangos de operación donde se encuentra efectivamente un mejoramiento en el costo de producción. Este resultado permite evidenciar el cumplimiento de uno de los principales objetos de estudio de este proyecto.

Tabla 5. Resultado análisis optimización con función Optimize en Python

ANÁLISIS REGRESIÓN LINEAL MÚLTIPLE DE PYTHON			
VARIABLES	COEFICIENTES	VALORES (MEAN)	VALORES ÓPTIMOS
Temperatura Entrada	1005,42	430	410
Temperatura_Yankee	-1402,81	115	117
Presion_diferencial_1	-465,24	70	80
Presion_diferencial_2	170,92	17,5	10
Presion_diferencial_3	-2633,48	15	20
Velocidad_Motor_1	-313,61	1550	1570
Velocidad_Motor_2	283,61	1515	1500
Velocidad_Motor_3	-4,91	1457,5	1645
Velocidad_Motor_4	-20,43	2870	3000
Velocidad_Motor_5	2,68	2730	2670
Velocidad_Motor_6	-167,88	920	925
Velocidad_Yankee_m/min	91,89	1541	1520
COSTO PRODUCTO POR HORA		\$ 961.746,36	\$ 911.879,34

Fuente: Completamente generada por el autor, 2023

Para el cumplimiento del otro importante resultado en la propuesta, se continuó en el camino de obtener predicción basada en redes neuronales artificiales, del costo resultante de producir ante condiciones que puedan ser diferentes a las de los resultados obtenidos por optimización.

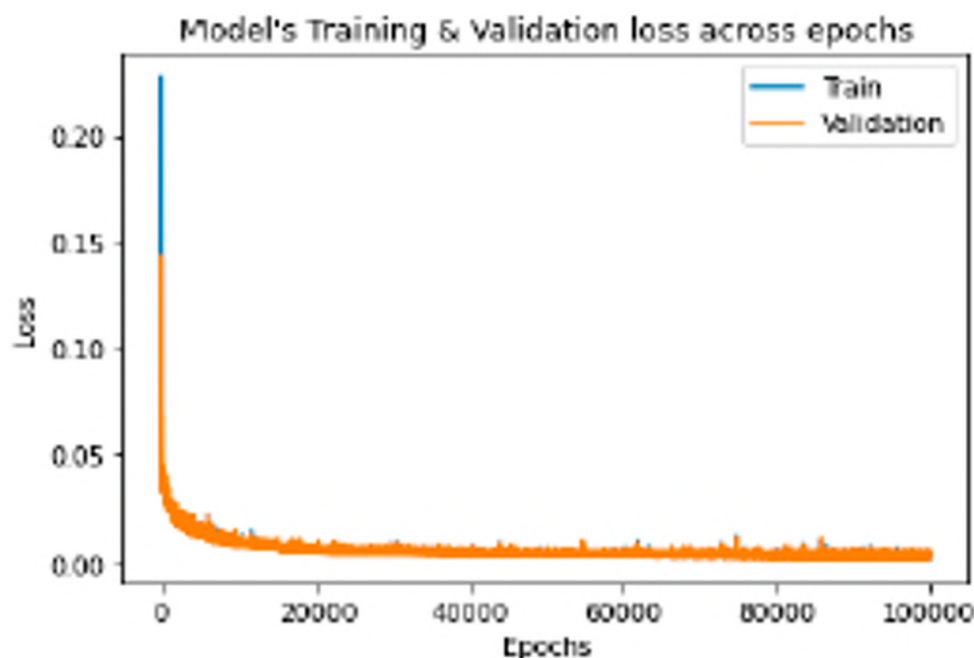
Se estableció como un valor base de aceptación del desempeño de la red neuronal, un Score superior al 90% en validación.

Una vez entrenado y validado se guardó el modelo de redes neuronales para calcular en tiempo real el costo, teniendo ya definido los límites operativos y conociendo las variables óptimas del proceso; se reitera que, si por razones de distinta índole en el proceso no se puede mantener al mínimo costo, puede ser debido a que gerencia necesita la mayor producción posible con lo que se generan cambios en los valores de trabajo que son distintos a los valores óptimos para mínimo costo en producción.

El score mínimo de 90% es solicitado por los expertos del proceso para garantizar mayor confiabilidad del resultado, adicionalmente en la literatura investigaba las redes neuronales llegan a este valor en la mayoría de los casos con un correcto entrenamiento.

La Figura 20, muestra gráficamente el mejor resultado obtenido donde se evidencia el ajuste del entrenamiento vs la validación y sigue la misma tendencia con pocas desviaciones que corresponden solo al 5%, cabe resaltar que, si bien se muestran solo una simulación, fueron cerca de 80 simulaciones las que se corrieron para obtener el resultado del entrenamiento de este algoritmo como lo es el R^2 que supero la condición establecida llegando al 95% de acierto en predicción.

Figura 20. Mejor resultado obtenido del entrenamiento de la Red Neuronal



Fuente: Completamente generada por el autor, 2023

4. DISCUSIÓN Y CONCLUSIONES

La experiencia de expertos en el proceso fue necesaria para definir los rangos operativos de la producción de papel, permitiendo el entendimiento adecuado del proceso y un preprocesamiento más sencillo, ya que la información fue previamente analizada.

Para este trabajo el método estadístico de Regresión Lineal múltiple permitió establecer el mínimo costo en la fabricación de papel doble hoja determinando los rangos operativos para la operación.

Aplicando la función Optimaze de Python se logró encontrar un costo operativo mínimo de \$911.879,34 que, comparado con la predicción de valores promedio \$961.746,36, obteniendo como resultado un ahorro mensual de \$38.337.225,64 COP/Mes, porcentualmente corresponde al 5,18%.

La Red Neuronal Artificial utilizada alcanzó un muy buen desempeño con un score de R^2 95% entregando resultados confiables y generando el modelo de apoyo para predecir el costo según las variables operativas del proceso.



REFERENCIAS

- IBM. (2023). Retrieved October 21, 2023 ¿Qué es el aprendizaje supervisado? | from <https://www.ibm.com/es-es/topics/supervised-learning>
- Aldrich, C., & Auret, L. (2013). *Unsupervised process monitoring and fault diagnosis with machine learning methods*. Springer. <https://link.springer.com/book/10.1007%2F978-1-4471-5185-2>
- Bottou, L., Curtis, F. E., & Nocedal, J. (2018). Optimization methods for large-scale machine learning. *SIAM Review*, 60(2), 223–311. <https://arxiv.org/pdf/1606.04838.pdf>
- Caluña, C., & Patricio, E. (2018). Incremento de la productividad en el molino mp5 de fabricación de papel tissue, de la empresa Familia Sancela del Ecuador SA, mediante la teoría de las restricciones “TOC”. Quito, 2018. <http://bibdigital.epn.edu.ec/bitstream/15000/19203/1/CD-8577.pdf>
- Câmara, M. M., Soares, R. M., Feital, T., Anzai, T. K., Diehl, F. C., Thompson, P. H., & Pinto, J. C. (2017). Numerical Aspects of Data Reconciliation in Industrial Applications. *Processes*, 5(4), 56. https://scholar.google.com/scholar?hl=es&as_sdt=0%2C5&q=Numerical+Aspects+of+Data+Reconciliation+in+Industrial+Applications&btnG=
- Edwards, P. J., Murray, A. F., Papadopoulos, G., Wallace, A. R., Barnard, J., & Smith, G. (1999). The application of neural networks to the papermaking industry. *IEEE Transactions on Neural Networks*, 10(6), 1456–1464.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM*, 39(11), 27–34. [http://shawndra.pbworks.com/f/The KDD process for extracting useful knowledge from volumes of data.pdf](http://shawndra.pbworks.com/f/The+KDD+process+for+extracting+useful+knowledge+from+volumes+of+data.pdf)
- Fleiter, T., Fehrenbach, D., Worrell, E., & Eichhammer, W. (2012). Energy efficiency in the German pulp and paper industry—A model-based assessment of saving potentials. *Energy*, 40(1), 84–99. http://kchbi.chtf.stuba.sk/upload_new/file/Miro/Proc+problemy+odovzdane+zadania/Molnár/fleiter2012.pdf
- Ge, Z., Song, Z., Ding, S. X., & Huang, B. (2017). Data mining and analytics in the process industry: the role of machine learning. *IEEE Access*, 5, 20590–20616. <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8051033>
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier. <http://rizalespe.lecture.ub.ac.id/files/2015/10/DM-04-1-Klasifikasi.pdf>
- Haque, M. T., & Kashtiban, A. M. (2000). Application of neural networks in power systems; A review. *Power*, 2005.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). Unsupervised learning. In *The elements of statistical learning* (pp. 485–585). Springer.
- Hurtado, S. M., & Aguado, J. G. (2005). Predicción de demanda de energía en Colombia mediante un sistema de inferencia difuso neuronal. *Energética*, 33, 15–24.
- Jolliffe, I. (2011). Principal component analysis. In *International encyclopedia of statistical science* (pp. 1094–1096). Springer.
- Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons. <http://dspace.fue.edu.eg/xmlui/bitstream/handle/123456789/3583/8857.pdf?sequence=1>
- Laurijssen, J., Frans, J., Worrell, E., & Faaij, A. (2010). Optimizing the energy efficiency of conventional multi-cylinder dryers in the paper industry. *Energy*, 35(9), 3738–3750. http://kchbi.chtf.stuba.sk/upload_new/file/Miro/Proc+problemy+odovzdane+zadania/Čopák/Optimizing+the+energy+efficiency+of+conventional+multi-cylinder+dryers+in+the+paper+industry_uz.pdf



REFERENCIAS

- Maarouf, M. (2016). Aplicación de redes neuronales para la resolución de problemas complejos de confiabilidad y riesgo. https://scholar.google.es/scholar?hl=es&as_sdt=0%2C33&q=Aplicación+de+Redes+Neuronales+para+la+Resolución+de+Problemas+Complejos+en+Confiabilidad+y+Riesgo&btnG=
- Maisueche Cuadrado, A. (2019). Utilización del Machine Learning en la industria 4.0.
- Maldonado, H. O. S., & Acevedo, W. M. V. (n.d.). artificial intelligence in forecasting demands for electricity: an application in optimization of energy resources inteligencia artificial en pronóstico de demanda de energía eléctrica: una aplicación en optimización de recursos energéticos. https://scholar.google.es/scholar?hl=es&as_sdt=0%2C33&q=Pronóstico+de+demanda+de+energía+eléctrica+en+la+región+Antioquia-Choco+utilizando+Redes+Neuronales+Artificiales.&btnG=
- Marques, M. P. (2014). Minería de Datos. La Metodología Cris-DM de Ibm. El Lenguaje Ciem E IBM SPSS Modeler: CreateSpace Independent Publishing Platform.
- Meireles, M. R. G., Almeida, P. E. M., & Simões, M. G. (2003). A comprehensive review for industrial applicability of artificial neural networks. *IEEE Transactions on Industrial Electronics*, 50(3), 585–601.
- Menacho Chiok, C. H. (2014). Modelos de regresión lineal con redes neuronales. *Anales Científicos*, ISSN-e 2519-7398, Vol. 75, No. 2, 2014 (Ejemplar Dedicado a: Julio a diciembre), Págs. 253-260, 75(2), 253–260. <https://dialnet.unirioja.es/servlet/articulo?codigo=7025156&info=resumen&idioma=SPA>
- Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.
- Moolayil, J., Moolayil, J., & John, S. (2019). *Learn Keras for Deep Neural Networks*. Springer.
- Nemeth, M., & Michalconok, G. (2017). The initial analysis of failures emerging in production process for further data mining analysis. *Process Control (PC)*, 2017 21st International Conference On, 210–215. https://scholar.google.com/scholar?hl=es&as_sdt=0%2C5&q=The+initial+analysis+of+failures+emerging+in+production+process+for+further+data+mining+analysis&btnG=#d=gs_cit&p=&u=%2Fscholar%3Fq%3Dinfo%3A-LcvZ5PE9YsJ%3AAscholar.google.com%2F%26output%3Dcite%26sci
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., & Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825–2830.
- Proceso de Fabricación del Papel. (2020). Retrieved September 8, 2020, from <http://emprendejc.blogspot.com/>
- Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier. https://books.google.com.co/books?hl=es&lr=&id=b3ujBQAAQBAJ&oi=fnd&pg=PP1&dq=related:73uIXCf8NGkJ:scholar.google.com/&ots=sQ5vYLFtG4&sig=u1sH3X_NCU450tyJBGuN0etvyp0#v=onepage&q&f=false
- Rojo Abuín, J. M. (2007). Regresión lineal múltiple.
- Shang, C., Huang, X., & You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers & Chemical Engineering*, 106, 464–479.
- Simon, P. (2013). *Too big to ignore: the business case for big data* (Vol. 72). John Wiley & Sons.
- Sivill, L., Ahtila, P., & Taimisto, M. (2005). Thermodynamic simulation of dryer section heat recovery in paper machines. *Applied Thermal Engineering*, 25(8–9), 1273–1292. <http://lib.tkk.fi/Diss/2011/isbn9789526042398/article7.pdf>



REFERENCIAS

- Tabares-Ospina, H., & Hernández-Riveros, J. (2009). *Pronósticos Demanda de Energía y Potencia a Corto Plazo. Caso de Estudio: Unidad de Control de Pronóstico Noroeste (Área Antioquia-Chocó)*. https://scholar.google.es/scholar?hl=es&as_sdt=0%2C33&q=Pronóstico+de+demand+de+energía+elétrica+en+la+región+Antioquia-Choco+utilizando+Redes+Neuronales+Artificiales.&btnG=
- Trokhan, P. D. (1987). *Tissue paper*. Google Patents. <https://patentimages.storage.googleapis.com/e1/8b/d8/3646c26538f81f/US4637859.pdf>
- Valmet. (2013). *Fabrics - History and Forming Fabrics*.
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann. <https://books.google.com.co/books?hl=es&lr=&id=1SyICgAAQBAJ&oi=fnd&pg=PP1&dq=related:73ulXCf8NGkJ:scholar.google.com/&ots=8IDRxdIDAa&sig=bfwbNcLtxVVB11yKtJn5s4oilWY#v=onepage&q&f=false>
- Zhao, J., Wang, W., & Sheng, C. (2018). *Data-Driven Prediction for Industrial Processes and Their Applications*. Springer. https://books.google.com.co/books?hl=es&lr=&id=vF5qDwAAQBAJ&oi=fnd&pg=PR5&ots=A-8CII70Nw&sig=29_Ce1tgtaRRyDlsoQY1IWj2q00&redir_esc=y#v=onepage&q&f=false
- Zhu, J., Ge, Z., & Song, Z. (2017). *Distributed parallel PCA for modeling and monitoring of large-scale plant-wide processes with big data*. *IEEE Transactions on Industrial Informatics*, 13(4), 1877–1885. <https://epm.elogim.com:2079/document/7835187?denied=>