



Revista de Investigación en Tecnologías de la Información

ISSN: 2387-0893

editorial@riti.es

Universitat Politècnica de Catalunya

España

Báez Castillo, Pablo; Reyes Hernández, Luis Ángel; Olivares Zepahua, Beatriz Alejandra; López Martínez, Ignacio; Córdova Ovando, María Isabel
Predicción de lectura en instrumento patrón para una empresa de metrología a través de un *dashboard* empleando algoritmos de regresión

Revista de Investigación en Tecnologías de la Información,
vol. 12, núm. 27, Esp., 2024, mayo-octubre, pp. 26-39

Universitat Politècnica de Catalunya

España

DOI: <https://doi.org/10.36825/RITI.12.27.004>

- Número completo
- Más información del artículo
- Página de la revista en portal.amelica.org





Predicción de lectura en instrumento patrón para una empresa de metrología a través de un *dashboard* empleando algoritmos de regresión

Reading prediction in a standard instrument for a metrology company through a dashboard using regression algorithms

Pablo Báez Castillo

Instituto Tecnológico de Orizaba, Veracruz, México
M17010193@orizaba.tecnm.mx
ORCID: 0009-0006-3935-625X

Luis Ángel Reyes Hernández

Instituto Tecnológico de Orizaba, Veracruz, México
luis.rh@orizaba.tecnm.mx
ORCID: 0000-0002-0059-2996

Beatriz Alejandra Olivares Zepahua

Instituto Tecnológico de Orizaba, Veracruz, México
beatriz.oz@orizaba.tecnm.mx
ORCID: 0000-0003-2799-0887

Ignacio López Martínez

Instituto Tecnológico de Orizaba, Veracruz, México
Ignacio.lm@orizaba.tecnm.mx
ORCID: 0000-0001-7786-6521

María Isabel Córdova Ovando

Metrología y Suministros de Veracruz, S.A. de C.V.
calidad@msvcompany.com.mx

doi: <https://doi.org/10.36825/RITI.12.27.004>

Recibido: Junio 22, 2024

Aceptado: Agosto 17, 2024

Resumen: Actualmente, la metrología ha empleado técnicas avanzadas como la inteligencia artificial, el internet de las cosas y el análisis de datos, todo con el objetivo de medir de manera precisa y exacta diferentes magnitudes físicas. Todo esto se le conoce hoy en día como metrología 4.0, que tiene como objetivo principal asegurar la fiabilidad y la precisión de mediciones en campos como la industria. En este contexto, se propone una aplicación web para una empresa que se dedica a la calibración de instrumentos, en la cual se apliquen algoritmos de regresión para la predicción en sus instrumentos de calibración. Esto no solo permitirá ofrecer resultados de mejor calidad y más confiables a los clientes, sino también permitirá optimizar los procesos de la empresa y mejorar la eficiencia en la toma de decisiones relacionadas con la calibración de los instrumentos. Se espera que, mediante el uso de

algoritmos de regresión, la empresa prediga cuándo el instrumento patrón va a dar malas lecturas. Esto permitirá enviar el instrumento a calibrar con anticipación, evitando errores y garantizando la precisión en las mediciones. Aunque aún no se tienen resultados concretos, se anticipa una mejora significativa en la precisión y fiabilidad de las predicciones, así como en la eficiencia operativa de la empresa.

Palabras clave: *Aplicación Web, Predicción, Metrología, Algoritmos de Regresión.*

Abstract: Nowadays, metrology has employed advanced techniques such as artificial intelligence, the internet of things and data analysis, all with the aim of accurately and precisely measuring different physical magnitudes. All this is known today as metrology 4.0, which has as its main objective to ensure the reliability and accuracy of measurements in fields such as industry. In this context, a web application is proposed for a company dedicated to the calibration of instruments, in which regression algorithms are applied for prediction in their calibration instruments. This will not only provide better quality and more reliable results to customers, but will also optimize the company's processes and improve the efficiency of decision making related to instrument calibration. It is expected that, through the use of regression algorithms, the company will predict when the standard instrument will give bad readings. This will make it possible to send the instrument for calibration in advance, avoiding errors and guaranteeing measurement accuracy. Although concrete results are not yet available, a significant improvement in the accuracy and reliability of predictions is anticipated, as well as in the company's operational efficiency.

Keywords: *Web Application, Prediction, Metrology, Regression Algorithms.*

1. Introducción

Actualmente, la empresa Metrología y Suministros de Veracruz S.A. de C.V, ubicada en la zona centro del estado Veracruz, más específicamente en la ciudad de Río Blanco, se dedica a la calibración de instrumentos y cuenta con una gran cantidad de clientes quienes envían sus instrumentos con ellos para ser calibrados. El proceso consiste en recibir los instrumentos que los clientes envían a calibrar y dependiendo del tipo de instrumento, la empresa los canaliza al laboratorio correspondiente para llevar a cabo el proceso de calibración.

El proceso de calibración consiste en calcular la incertidumbre de cada instrumento. La estimación de incertidumbre [1], se refiere al proceso de cuantificar la falta de certeza asociada con una medición en el contexto de la metrología. La estimación de la incertidumbre se basa en la evaluación de todos los factores que contribuyen a la falta de precisión en una medición, incluyendo errores sistemáticos y aleatorios.

El primer paso del proceso de calibración es definir los puntos de calibración, los cuales son indicados por el cliente, de no ser así, se toman los puntos de calibración cada 10, 50 y 90% de la escala partiendo de su límite menor. Después, se selecciona el instrumento patrón y el medio de generación de temperatura en el que se sumerge el instrumento. El siguiente paso es preparar el instrumento a calibrar y el instrumento patrón a las mismas condiciones (condiciones de baño, horno u otro medio que reproduce, mide o simula una temperatura especificada). Luego se toman cuatro lecturas de temperatura en 10%, 50% y 90% de la escala partiendo de su límite menor o de acuerdo a los requerimientos del cliente. Son 4 lecturas del instrumento a calibrar y 4 lecturas del instrumento patrón y se anotan en la toma de datos para la calibración del instrumento. Y, por último, se apagan los equipos utilizados.

Hoy en día, la empresa Metrología y Suministros de Veracruz S.A de C.V. realiza la captura de las lecturas en documentos de Excel, en los cuales, también se encuentran las fórmulas utilizadas para calcular la incertidumbre de cada instrumento. A pesar de que el proceso funciona, la empresa tiene la necesidad de mejorar los tiempos para realizar más calibraciones en menos tiempo.

En este contexto, se propone el desarrollo de una aplicación web que permita a los usuarios introducir las lecturas de los instrumentos y realizar las operaciones correspondientes para la estimación de incertidumbre de manera automatizada. Además, se busca implementar un *dashboard* para mejorar la gestión y el control de operaciones. Según el trabajo de Cifuentes Linares [2], los *dashboards* son eficaces, concisos y precisos en la gestión de datos, lo cual ayudará a la empresa a tener un mejor control con sus calibraciones y sus clientes.

Asimismo, el objetivo de la investigación es la elección de un algoritmo de regresión en el *dashboard* para predecir las lecturas del instrumento patrón en futuras calibraciones y mejorar la precisión en los resultados. De la

Fuente Ruiz [3] menciona que la metrología 4.0, como aplicación de la industria 4.0 en el campo de la metrología, utiliza tecnologías avanzadas como sensores inteligentes, internet de las cosas (IoT), análisis de datos y *big data* para mejorar la precisión y la eficiencia de las mediciones, así como para optimizar los procesos de producción y reducir costos.

2. Estado del arte

En esta sección se analiza la literatura más relacionada con la metrología, el desarrollo del *dashboard* y los algoritmos de regresión.

En primer lugar, en [4] se abordó la enseñanza de la metrología, destacando su importancia en el desarrollo de capacidades de medición y su papel crucial en la innovación y la garantía de calidad de productos. Entre las ventajas, se mencionó que la metrología es fundamental para el desarrollo económico y social, así como para la competitividad empresarial y la protección del consumidor. Sin embargo, también se identificaron debilidades, como la falta de visibilidad de la metrología como disciplina y la confusión con otras áreas, lo que ha llevado a una baja demanda en programas de formación específicos.

Por otro lado, en [3] se mencionó la integración de tecnologías avanzadas en la medición y control de variables climáticas, destacando las ventajas de la Industria 4.0, como la personalización de productos, la optimización de la calidad y la mejora en la toma de decisiones a través del análisis de datos en tiempo real. Sin embargo, también se identificaron desventajas, como la complejidad en la gestión de grandes volúmenes de datos y la necesidad de herramientas específicas para su análisis, lo que puede representar una debilidad en la implementación de estas tecnologías.

A pesar de estos desafíos, el nicho de oportunidad radica en la creciente demanda de soluciones metroológicas precisas y eficientes, impulsadas por la interconexión de dispositivos y el uso de inteligencia artificial, lo que sugiere un potencial significativo para el desarrollo futuro en este campo. Los *dashboards* son herramientas para la gestión de datos y en [5] se mencionó la implementación de un *dashboard* basado en minería de datos para mejorar la gestión de contratos en el sector público en Colombia. Entre las ventajas, se destacó la capacidad de identificar patrones y comportamientos anómalos en los procesos de contratación. Sin embargo, se identificaron debilidades en la calidad de los datos, lo que representa un nicho de oportunidad para mejorar la calidad de los mismos y desarrollar herramientas más accesibles y especializadas en el análisis de variables categóricas.

Así como también en [6] se presentó el diseño e implementación de *dashboards* para la visualización de clústeres. Se describió el proceso KDD (*Knowledge Discovery in Databases*) y su relación con la minería de datos. También se explicó la minería de datos y se detallaron los modelos de minería que se utilizaron para la clasificación no supervisada en *clustering*. Con este trabajo se puede, nuevamente, confirmar la importancia y lo útil que es un *dashboard* para las organizaciones, ya que les permite tener datos importantes a la vista y entenderlos de manera fácil. Por otra parte, en [7] se mencionó el uso de algoritmos de aprendizaje automático en el análisis y predicción de datos, destacando su capacidad para identificar tendencias y comportamientos futuros a partir de grandes volúmenes de información. Entre las ventajas se mencionó la precisión y estabilidad de modelos como los árboles de decisión y las redes neuronales, que permiten representar relaciones complejas y realizar predicciones efectivas. En [8] se abordó la aplicación de algoritmos de regresión lineal, tanto simple como múltiple, para analizar la relación entre la temperatura y la humedad en un *datacenter*, con el objetivo de optimizar la ventilación y prolongar la vida útil de los servidores. Entre las ventajas se destacó la capacidad de modelar y predecir condiciones climáticas que afectan el rendimiento de los sistemas informáticos. Sin embargo, una debilidad que se identificó es que la regresión lineal puede no ser adecuada para todos los escenarios, especialmente en la temperatura, esto representa una oportunidad para investigar y desarrollar modelos más robustos que puedan aplicarse en diversos contextos donde la ventilación también es crucial.

3. Materiales y métodos

A continuación, se presentan los datos más importantes que fueron extraídos de los archivos de instrumentos calibrados por la empresa Metrología y Suministros de Veracruz S.A de C.V. Estos datos fueron seleccionados por el autor, en colaboración con un experto en algoritmos de inteligencia artificial, con el objetivo de tener información relevante sobre los instrumentos para aplicar los algoritmos de inteligencia artificial. La identificación

de estos datos se basó en su relevancia para el análisis y aplicación de algoritmos de regresión. Ya que el objetivo es elegir el algoritmo de regresión adecuado para predecir las lecturas del instrumento patrón.

Asimismo, se presentan los algoritmos de regresión que se tomaron en cuenta para la elección del mejor de ellos y por último se muestran los *mockups* de la aplicación web.

3.1. Extracción de datos

La empresa Metrología y Suministros de Veracruz S.A de C.V no cuenta con los datos de sus calibraciones en una base de datos, actualmente tienen las calibraciones en documentos de Excel, es decir, un documento por cada calibración. En el cual se encuentran las lecturas del equipo patrón. La información de las lecturas se encuentran las tomas de datos que corresponden a los documentos de Excel, las cuales se muestran en las Figuras 1 y 2.

Mediciones									
Temperatura	10.0			20.0			25.0		
Lectura	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Instrumento
1	10.003	10.0	20.005	19.7	25.008	24.7			
2	10.003	10.0	20.005	19.7	25.008	24.7			
3	10.003	10.0	20.005	19.7	25.008	24.7			
4	10.003	10.0	20.005	19.7	25.008	24.7			
M. Térmico:	MSV/EAT-03 (+)			MSV/EAT-03 (+)			MSV/EAT-03 (+)		
Temperatura	10.0			20.0			25.0		
Lectura	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Instrumento
1									
2									
3									
4									
M. Térmico:									
Temperatura	10.0			20.0			25.0		
Lectura	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Instrumento
1									
2									
3									
4									
M. Térmico:									
Temperatura	10.0			20.0			25.0		
Lectura	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Instrumento
1									
2									
3									
4									
M. Térmico:									

Figura 1. Toma de datos de termómetros de lectura directa (TLD).

Mediciones									
Temperatura	0.0			2.0			4.0		
Lectura	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Instrumento
1	0.003	0.0	2.014	2.0	4.005	4.0	6.021	6.0	
2	0.003	0.0	2.014	2.0	4.005	4.0	6.021	6.0	
3	0.003	0.0	2.014	2.0	4.005	4.0	6.021	6.0	
4	0.003	0.0	2.014	2.0	4.005	4.0	6.021	6.0	
M. Térmico:	MSV/EAT-03 (+)			MSV/EAT-03 (+)			MSV/EAT-03 (+)		
Temperatura	0.0			2.0			4.0		
Lectura	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Instrumento
1									
2									
3									
4									
M. Térmico:									
Temperatura	0.0			2.0			4.0		
Lectura	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Patrón	Instrumento	Instrumento
1									
2									
3									
4									
M. Térmico:									

Figura 2. Toma de datos de termómetros de líquido en vidrio (TLV).

En ambas figuras se encuentra la información de las mediciones, las cuales obtienen 4 lecturas del instrumento a calibrar y el instrumento patrón. Estas 4 lecturas se realizan por cada punto a calibrar.

A pesar de que las imágenes luzcan similares, no lo son. Es verdad que comparten los datos de las mediciones, pero en la Figura 2 se tiene información extra como las lecturas del termómetro auxiliar. Esto es importante mencionarlo para la extracción.

Para lograr extraer los datos numéricos de las lecturas de los instrumentos, se realizaron 2 *scripts* en Python. La razón de que fueron 2 *scripts* es porque, como se mencionó, al ser 2 tomas de datos diferentes, los datos se

encuentran en diferentes celdas. En la Figura 3 se muestra la selección de las celdas del documento de Excel de donde se van a extraer los datos. Y en la Figura 4 se muestra el resto del código del *script*.

```
import os
import xlwings as xw

# RUTA DE LA CARPETA CON LOS ARCHIVOS EXCEL
carpeta_excel = 'C:/Users/pablo/Documents/Maestria/Semestre 3/Tesis/Temperatura/Calibraciones/2024/TLD/'

# RUTA DEL EXCEL DESTINO
excel_destino = 'C:/Users/pablo/Documents/Maestria/Semestre 3/Tesis/Temperatura/Calibraciones/ConjuntoDatosTLD - copia - copia.xlsx'

# Lista de celdas
celdas = ['C11', 'G11', 'L11', 'H12', 'M12', 'R12', 'C13', 'P13', 'T13', 'C14', 'G51', 'G52', 'G53', 'P51', 'P52', 'P53', 'D17',
          'C19', 'E19', 'C20', 'E20', 'C21', 'E21', 'C22', 'E22', 'G17', 'G19', 'I19',
          'G20', 'I20', 'G21', 'I21', 'G22', 'I22', 'K17', 'K19', 'M19', 'K20', 'M20',
          'K21', 'M21', 'K22', 'M22', 'O17', 'O19', 'Q19', 'O20', 'Q20', 'O21', 'Q21',
          'Q22', 'Q25', 'C27', 'E27', 'C28', 'E28', 'C29', 'E29', 'C30', 'E30',
          'G25', 'G27', 'I27', 'G28', 'I28', 'G29', 'I29', 'G30', 'I30', 'K25', 'K27',
          'M27', 'K28', 'M28', 'K29', 'M29', 'K30', 'M30', 'O25', 'O27', 'Q27', 'Q28',
          'Q28', 'O29', 'Q29', 'O30', 'Q30', 'O33', 'C35', 'E35', 'C36', 'E36', 'C37',
          'E37', 'C38', 'E38', 'G33', 'G35', 'I35', 'G36', 'I36', 'G37', 'I37', 'G38',
          'I38', 'K33', 'K35', 'M35', 'K36', 'M36', 'K37', 'M37', 'K38', 'M38', 'O33',
          'O35', 'Q35', 'O36', 'Q36', 'O37', 'Q37', 'O38', 'Q38', 'O41', 'C43', 'E43',
          'C44', 'E44', 'C45', 'E45', 'C46', 'E46', 'G41', 'G43', 'I43', 'G44', 'I44',
          'G45', 'I45', 'G46', 'I46', 'K41', 'K43', 'M43', 'K44', 'M44', 'K45', 'M45',
          'M46', 'O41', 'O43', 'Q43', 'O44', 'Q44', 'O45', 'Q45', 'O46', 'Q46']
```

Figura 3. Selección de celdas del documento de TLD.

```
# Abrir la aplicación de Excel en segundo plano
app = xw.App(visible=False)

# Abrir el libro de Excel destino
wb_destino = app.books.open(excel_destino)

# Seleccionar la hoja de trabajo por su nombre
ws_destino = wb_destino.sheets['Hojal']

# Obtener la fila donde se agregarán los datos (fila siguiente a la última fila con datos)
fila_inicial = ws_destino.range('A' + str(ws_destino.cells.last_cell.row)).end('up').row + 1

# Recorrer todos los archivos en la carpeta
for archivo in os.listdir(carpeta_excel):
    if archivo.endswith('.xlsx'):
        ruta_archivo = os.path.join(carpeta_excel, archivo)

        try:
            # Abrir el libro de trabajo (workbook) del archivo de origen
            wb_origen = app.books.open(ruta_archivo)

            # Seleccionar la hoja de trabajo (worksheet) del archivo de origen por su nombre
            ws_origen = wb_origen.sheets['Toma de Datos TLV']

            # Obtener los datos de las celdas especificadas y agregarlos a la lista
            datos_a_agregar = [ws_origen.range(celda).value for celda in celdas]

            # Escribir los datos en el libro de Excel destino
            ws_destino.range(f'A{fila_inicial}').value = datos_a_agregar

            # Incrementar la fila inicial para la próxima inserción de datos
            fila_inicial += 1

            # Cerrar el libro de trabajo del archivo de origen
            wb_origen.close()

        except Exception as e:
            print(f"No se pudo acceder al archivo '{archivo}': {e}")

# Guardar los cambios en el libro de Excel destino
wb_destino.save()
wb_destino.close()
```

Figura 4. Extracción de datos de celdas.

3.2. Modelos de regresión

Los modelos de regresión que se investigaron fueron los siguientes:

1. Regresión lineal.

Es la técnica de modelado más utilizada, la cual tiene una conexión entre una variable dependiente (Y) y una variable independiente (X) [9]. Y la gráfica que mejor representa esta técnica se muestra en la Figura 5.

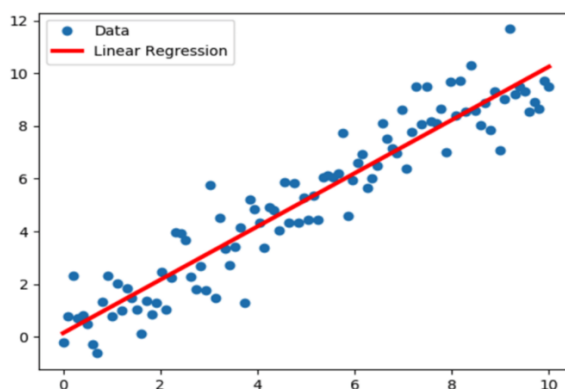


Figura 5. Regresión lineal.

2. Regresión polinómica.

La regresión polinómica es muy similar a la regresión lineal, ya que también existe relación entre variables dependientes e independientes [9]. Solamente que esta técnica se utiliza para representar una relación no lineal, así como se muestra en la Figura 6.

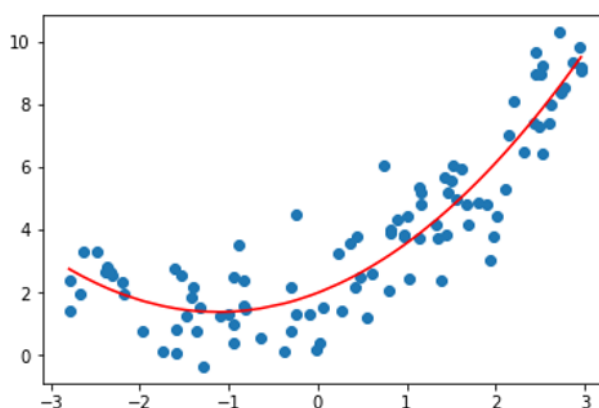


Figura 6. Regresión polinómica.

3. Árboles de regresión.

Los árboles de regresión son un método de aprendizaje supervisado utilizado para predecir valores continuos. Este algoritmo divide el espacio de características en subconjuntos más pequeños y, en cada división, se selecciona la característica y el punto de corte que minimizan el error de predicción. La estructura resultante es un árbol binario donde cada nodo interno representa una decisión basada en una característica y cada nodo terminal (hoja) representa una predicción del valor continuo. En la Figura 7 se observa la gráfica de esta técnica.

- a. Selección de características y punto de corte:
 - Se evalúan múltiples puntos de corte para cada característica del conjunto de datos.
 - Para cada posible punto de corte, se calcula una métrica de error, comúnmente el Error Cuadrático Medio (MSE, por sus siglas en inglés).
 - Se selecciona la característica y el punto de corte que minimicen el MSE, dividiendo así el conjunto de datos en dos subconjuntos.
- b. División recursiva:
 - El proceso de selección de característica y punto de corte se repite recursivamente para cada subconjunto resultante.
 - Esta recursión continúa hasta que se cumple un criterio de parada, como alcanzar una profundidad máxima del árbol o un número mínimo de datos por nodo.
- c. Nodos terminales:
 - Los nodos terminales del árbol, conocidos como hojas, contienen las predicciones finales.
 - La predicción en cada hoja es el valor promedio de los valores objetivos de los datos que llegan a esa hoja.

d. Ventajas:

- Simplicidad e Interpretabilidad: Los árboles de regresión son intuitivos y sus resultados son fácilmente interpretables.
- Manejo de Datos No Lineales: Pueden capturar relaciones no lineales entre las características y el valor objetivo sin necesidad de transformaciones complejas.

e. Desventajas:

- Propensión al Sobreajuste: Pueden sobre ajustarse a los datos de entrenamiento, especialmente si el árbol es muy profundo.
- Inestabilidad: Pequeñas variaciones en los datos pueden resultar en estructuras de árbol significativamente diferente [10].

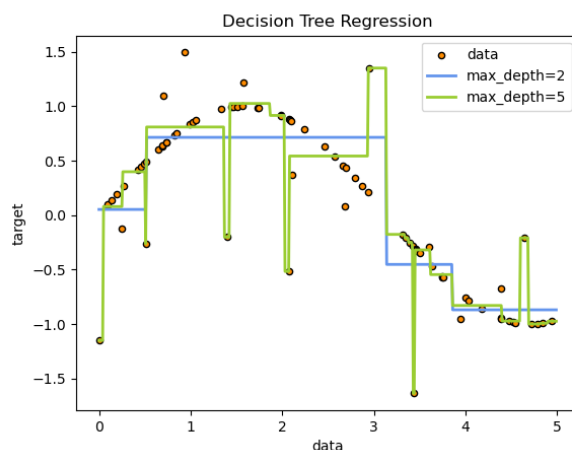


Figura 7. Árboles de regresión.

3.3. Mockups

A continuación, se presentan los *mockups* utilizados para representar cómo se verá el diseño final de la aplicación web. La toma de datos de termómetro de lectura directa (TLD), en la cual se solicita la información de las mediciones, es decir, 4 lecturas del instrumento patrón y las 4 lecturas del instrumento a calibrar. Posteriormente, se solicita ingresar las condiciones ambientales en las cuales se realizó la toma de datos (Figura 8).

Figura 8. Toma de datos TLD.

Por último, en la Figura 9, se muestra el diseño del *dashboard* para la aplicación web, en el cual se incluyen gráficas con datos importantes como el número total de instrumentos que hay sin calibrar, cuantos ya han sido calibrados, entre otros.

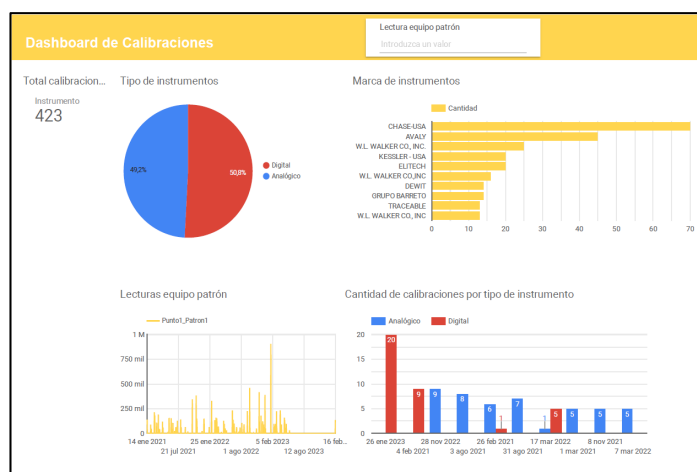


Figura 9. Dashboard para la aplicación web.

3.4. Patrón arquitectónico

En esta sección se presenta el patrón arquitectónico propuesto para el desarrollo de la aplicación web. La aplicación es diseñada empleando el patrón de diseño MVC (*Model-View-Controller*, Modelo-Vista-Contrólador). MVC es un patrón de diseño utilizado para crear aplicaciones web y es usado en casi todos los marcos de desarrollo web [11]. La arquitectura se muestra en la Figura 10.

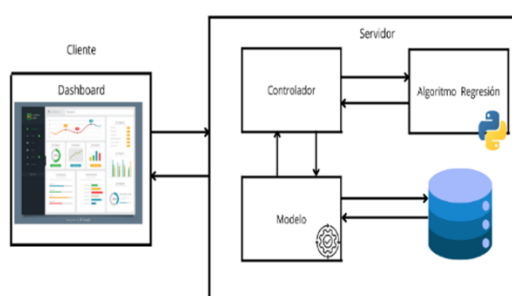


Figura 10. Patrón arquitectónico MVC propuesto para el desarrollo de la aplicación web.

En MVC, el modelo representa los datos y la lógica del negocio. Es el encargado de gestionar el acceso a la base de datos, realizar cálculos y procesamiento de la información. La vista es la interfaz del usuario que representa los datos al usuario y recibe las interacciones de este. Se encarga de mostrar la información de manera visual. El control es el intermediario entre el modelo y la vista. Responde a las acciones del usuario, actualiza el modelo según sea necesario y selecciona la vista que se debe mostrar [12].

A. Cliente

El término cliente se refiere a un componente en un sistema cliente-servidor que solicita servicios o recursos de un servidor. En el contexto de la programación cliente-servidor, un cliente es una aplicación que envía solicitudes a un servidor para acceder a datos, recursos o funcionalidades proporcionadas por el servidor [13].

B. Servidor

En el contexto de la programación cliente-servidor, un servidor es un componente que proporciona servicios, recursos o funcionalidades a los clientes que lo solicitan. Los servidores son programas o dispositivos que están diseñados para atender las peticiones de los clientes y proporcionarles los datos o servicios requeridos [13].

- **Modelo:** representa la lógica de la aplicación en distintas clases encargadas de estructurar e interactuar con la información de los instrumentos para utilizar el algoritmo de predicción.

- **Controlador:** es el componente que se encarga de la interacción entre el cliente y el servidor. Actúa como intermediario, recibiendo las solicitudes del cliente, procesándolas y enviándolas al servidor. En este caso va a recibir los datoos numéricos que el cliente introduzca para después utilizar el algoritmo de regresión.

4. Resultados

A continuación, se muestran los resultados de la extracción de datos de los documentos de Excel y del modelo del algoritmo de regresión. Una vez realizada la extracción de datos a través de los *scripts* en Python se obtuvieron dos conjuntos de datos que se muestran en las Figuras 11 y 12.

Modelo	Series	Div.Minim	Fecha_Cal	Resolucio	ID_instru	Numero	Liq_terminmetri	Punto_1	Punto1_P	Punto1_M	Punto1_N	Punto1_O	Punto1_P	Punto1_R	Punto1_S	grad_inm	Punto_2	Punto2_P	Punto2_M	Punto2_N	
ASTM-	889229	0.2	21/10/2021	0.1	889229	279 MERCURIO		-20	-19.216	-20	-19.216	-20	-19.216	-20	-19.216	-20	-20	-10	-10.004	-10.2	
	352586	0.5	11/10/2021	0.25	DIRE-IM-122	280 ALCOHOL		-20	-20.013	-19.5	-20.013	-19.5	-20.013	-19.5	-20.013	-19.5	-30	-15	-15.019	-14.5	
SIN SERIE	1	12/10/2021	0.5	DIRE-IM-53	287 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-28	2	2.014	2	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-54	288 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-34.5	2	2.014	2	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-55	289 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-57	2	2.014	2	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-56	290 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-50	2	2.014	2	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-57	291 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-26	55	55.019	55.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-58	292 MERCURIO	0	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	-39.5	55	55.019	55.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-59	293 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-36	35	35.018	35.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-60	294 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-28	35	35.018	35.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-61	295 MERCURIO	0	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	-50.5	35	35.018	35.5	
SIN SERIE	1	15/10/2021	0.5	DIRE-IM-62	296 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-49.5	35	35.018	35.5	
SIN SERIE	1	15/10/2021	0.5	DIRE-IM-63	297 MERCURIO	0	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	-30	68	68.01	68.5	
SIN SERIE	1	15/10/2021	0.5	DIRE-IM-64	298 MERCURIO	0	0.003	-1	0.003	-1	0.003	-1	0.003	-1	0.003	-1	-32	68	68.01	68.5	
ASTM-	25312	0.5	14/10/2021	0.25	25312	299 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-30	19	19.001	19
ASTM-	40668	0.2	14/10/2021	0.1	40668	300 MERCURIO	-20	-20.004	-19.9	-20.004	-19.9	-20.004	-19.9	-20.004	-19.9	-20.004	-19.9	-20	-10	-10.014	-10
ASTM-	14946	0.5	14/10/2021	0.25	14946	301 MERCURIO	-4	-4.007	-4	-4.007	-4	-4.007	-4	-4.007	-4	-4.007	-4	-5	15	15.028	15
ASTM-	2430	0.2	14/10/2021	0.1	2430	302 MERCURIO	75	75.008	75.2	75.008	75.2	75.008	75.2	75.008	75.2	75.008	75.2	62	90	90.012	90
ASTM-	1822	0.2	14/10/2021	0.1	1822	303 MERCURIO	100	100.005	100.2	100.005	100.2	100.005	100.2	100.005	100.2	100.005	100.2	90	115	115.021	115
SIN SERIE	1	15/10/2021	0.5	4	304 MERCURIO	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-35	0	0.003	0	
SIN SERIE	1	15/10/2021	0.5	11	305 MERCURIO	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-35	0	0.003	0	
ASTM-	816692	0.2	19/10/2021	0.1	816692	308 MERCURIO	20	20.009	20	20.009	20	20.009	20	20.009	20	20.009	20	-20	25	25.008	25
ASTM-	888596	0.1	19/10/2021	0.1	888596	309 MERCURIO	37.8	37.811	37.9	37.811	37.9	37.811	37.9	37.811	37.9	37.811	37.9	34	0	0.003	0
ASTM-	372	1	20/10/2021	0.5	SILM-TDES-0	310 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-2	10	10.005	10
ASTM-	2211 0138	0.5	16/11/2021	0.25	2211 0138	330 MERCURIO	-20	-20.011	-20	-20.011	-20	-20.011	-20	-20.011	-20	-20.011	-20	-29.5	0	0.003	0
ASTM-	A072244	1	17/11/2021	0.5	A072244	331 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-35	22	22.018	22
ASTM-	23478	0.5	17/11/2021	0.25	23478	332 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-22.5	19	19.025	19
ASTM-	706394	0.5	17/11/2021	0.25	706394	333 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-27.5	19	19.025	19.5
ASTM-	731244	0.5	17/11/2021	0.25	731244	334 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-24	19	19.025	19
ASTM-	19001	0.5	18/11/2021	0.25	19001	335 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-25.5	19	19.025	19
ASTM-	706178	0.5	18/11/2021	0.25	706178	336 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-30.5	19	19.025	19.5
ASTM-	7202	2	18/11/2021	1	7202	337 MERCURIO	100	100.036	102	100.036	102	100.036	102	100.036	102	100.036	102	42	200	200.023	200
ASTM-	643579	2	18/11/2021	1	643579	338 MERCURIO	100	100.036	104	100.036	104	100.036	104	100.036	104	100.036	104	24	200	200.023	204
ASTM-	40615	0.2	19/11/2021	0.1	40615	339 MERCURIO	-20	-19.043	-20	-19.043	-20	-19.043	-20	-19.043	-20	-19.043	-20	-20	-10	-9.507	-10
ASTM-	40638	0.2	19/11/2021	0.1	40638	340 MERCURIO	-20	-19.21	-20	-19.21	-20	-19.21	-20	-19.21	-20	-19.21	-20	-20	-10	-9.533	-10
ASTM-	15001	0.5	19/11/2021	0.25	15001	341 MERCURIO	-4	-4.018	-3.5	-4.018	-3.5	-4.018	-3.5	-4.018	-3.5	-4.018	-3.5	-5	15	15.026	15
SIN SERIE	1	22/11/2021	0.5	MSVC-01	342 ALCOHOL	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	40	30	30.03	30	

Figura 11. Conjunto de datos de TLD.

Modelo	Series	Div.Minisr	Fecha_Cal	Resolucio	ID_instru	Numero	Liq_terminmetri	Punto_1	Punto1_P	Punto1_M	Punto1_N	Punto1_O	Punto1_P	Punto1_R	Punto1_S	Punto1_T	grad_inm	Punto_2	Punto2_P	Punto2_M	Punto2_N	
ASTM-	889229	0.2	21/10/2021	0.1	889229	279 MERCURIO		-20	-19.216	-20	-19.216	-20	-19.216	-20	-19.216	-20	-20	-10	-10.004	-10.2	-10.2	
ASTM-	352586	0.5	11/10/2021	0.25	DIRE-IM-122	280 ALCOHOL		-20	-20.013	-19.5	-20.013	-19.5	-20.013	-19.5	-20.013	-19.5	-30	-15	-15.019	-14.5	-14.5	
SIN SERIE	1	12/10/2021	0.5	DIRE-IM-53	287 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-28	2	2.014	2	2.014	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-54	288 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-34.5	2	2.014	2	2.014	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-55	289 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-57	2	2.014	2	2.014	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-56	290 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-50	2	2.014	2	2.014	
SIN SERIE	1	13/10/2021	0.5	DIRE-IM-57	291 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-26	55	55.019	55.5	55.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-58	292 MERCURIO	0	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	-39.5	55	55.019	55.5	55.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-59	293 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-36	35	35.018	35.5	35.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-60	294 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-28	35	35.018	35.5	35.5	
SIN SERIE	1	14/10/2021	0.5	DIRE-IM-61	295 MERCURIO	0	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	-50.5	35	35.018	35.5	35.5	
SIN SERIE	1	15/10/2021	0.5	DIRE-IM-62	296 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-49.5	35	35.018	35.5	35.5	
SIN SERIE	1	15/10/2021	0.5	DIRE-IM-63	297 MERCURIO	0	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	0.003	0.5	-30	68	68.01	68.5	68.5	
SIN SERIE	1	15/10/2021	0.5	DIRE-IM-64	298 MERCURIO	0	0.003	-1	0.003	-1	0.003	-1	0.003	-1	0.003	-1	-32	68	68.01	68.5	68.5	
ASTM-	25312	0.5	14/10/2021	0.25	25312	299 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-30	19	19.001	19	19.001
ASTM-	40668	0.2	14/10/2021	0.1	40668	300 MERCURIO	-20	-20.004	-19.9	-20.004	-19.9	-20.004	-19.9	-20.004	-19.9	-20	-10	-10.014	-10	-10.014	-10.014	
ASTM-	14946	0.5	14/10/2021	0.25	14946	301 MERCURIO	-4	-4.007	-4	-4.007	-4	-4.007	-4	-4.007	-4	-5	15	15.028	15	15.028	15.028	
ASTM-	2430	0.2	14/10/2021	0.1	2430	302 MERCURIO	75	75.008	75.2	75.008	75.2	75.008	75.2	75.008	75.2	75.008	75.2	62	90	90.012	90.2	90.2
ASTM-	1822	0.2	14/10/2021	0.1	1822	303 MERCURIO	100	100.005	100.2	100.005	100.2	100.005	100.2	100.005	100.2	100.005	100.2	90	115	115.021	115.2	115.2
SIN SERIE	1	15/10/2021	0.5	11	304 MERCURIO	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-35	0	0.003	0	0.003	
SIN SERIE	1	15/10/2021	0.5	11	305 MERCURIO	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-10.014	-10	-35	0	0.003	0	0.003	
ASTM-	816692	0.2	19/10/2021	0.1	816692	308 MERCURIO	20	20.009	20	20.009	20	20.009	20	20.009	20	20.009	20	-20	25	25.008	25	25.008
ASTM-	888596	0.1	19/10/2021	0.1	888596	309 MERCURIO	37.8	37.811	37.9	37.811	37.9	37.811	37.9	37.811	37.9	37.811	37.9	34	20	20.003	20	20.003
ASTM-	22211038	0.1	20/10/2021	0.5	SIAM-T065	310 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-2	10	10.003	10	10.003
ASTM-	22211038	0.5	16/11/2021	0.25	22211038	330 MERCURIO	-20	-20.011	-20	-20.011	-20	-20.011	-20	-20.011	-20	-20.011	-20	-29.5	0	0.003	0	0.003
ASTM-	A07224A	1	17/11/2021	0.5	A07224A	331 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-35	22	22.018	22	22.018
ASTM-	23478	0.5	17/11/2021	0.25	23478	332 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-22.5	19	19.025	19	19.025
ASTM-	70614	0.5	17/11/2021	0.25	70614	333 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-20	19	19.025	19	19.025
ASTM-	731244	0.5	17/11/2021	0.25	731244	334 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-24	19	19.025	19	19.025
ASTM-	19001	0.5	18/11/2021	0.25	19001	335 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-25.5	19	19.025	19	19.025
ASTM-	706178	0.5	18/11/2021	0.25	706178	336 MERCURIO	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	-30.5	19	19.025	19	19.025
ASTM-	202	2	18/11/2021	1	202	337 MERCURIO	100	100.036	104	100.036	104	100.036	104	100.036	104	100.036	104	200	200.023	200	200.023	
ASTM-	643579	2	18/11/2021	1	643579	338 MERCURIO	100	100.036	104	100.036	104	100.036	104	100.036	104	100.036	104	24	200	200.023	200	200.023
ASTM-	406615	0.2	19/11/2021	0.1	406615	339 MERCURIO	-20	-19.043	-20	-19.043	-20	-19.043	-20	-19.043	-20	-19.043	-20	-10	-10.507	-10.5	-10.5	
ASTM-	406838	0.2	19/11/2021	0.1	406838	340 MERCURIO	-20	-19.211	-20	-19.211	-20	-19.211	-20	-19.211	-20	-19.211	-20	-10	-10.503	-10.5	-10.5	
ASTM-	15001	0.5	22/11/2021	0.25	15001	341 MERCURIO	-10	-10.018	-10	-10.018	-10	-10.018	-10	-10.018	-10	-10.018	-10	-15	19	19.025	19	19.025
SIN SERIE	1	22/11/2021	0.5	MSV-CP	342 ALCOHOL	0	0.003	0	0.003	0	0.003	0	0.003	0	0.003	0	0	50	50.003	50	50.003	

F	G	H	I	J	K	L	M
Fecha_Calibraci	Resoluci	ID_instrumen	Clase_exactitud	Numero_instrumen	Punto_1	Punto1_Patron	Diferencial_P
25/07/2022	0.5	30	0.5	302	15	15.008	6.4E-05
10/06/2022	0.5	1	0.5	233	15	15.01	1E-04
10/06/2022	0.5	2	0.5	234	15	15.01	1E-04
10/06/2022	0.5	3	0.5	235	15	15.01	1E-04
10/06/2022	0.5	4	0.5	236	15	15.01	1E-04
01/11/2022	0.5 SIN ID		0.5	396	10	10.024	0.000576
10/06/2022	0.5	5	0.5	237	15	15.01	1E-04
01/11/2022	0.5 SIN ID		0.5	395	10	10.024	0.000576
25/07/2022	0.5	2065	0.5	300	15	15.008	6.4E-05
31/03/2023	0.5	2093	0.5	117	15	15.01	1E-04
01/11/2022	0.5 SIN ID		0.5	397	10	10.024	0.000576
01/11/2022	0.5 SIN ID		0.5	399	10	10.024	0.000576
01/11/2022	0.5 SIN ID		0.5	398	10	10.024	0.000576
25/07/2022	0.5	2154	0.5	301	15	15.008	6.4E-05
31/03/2023	0.5	2734	0.5	116	15	15.01	1E-04
25/07/2022	0.5	2950	0.5	299	15	15.008	6.4E-05
12/09/2022	0.1 FARMACIACMT C		0.2	365	2	2.013	0.000169
28/11/2022	0.5 BX-LF-20-02		0.5	427	20	20.014	0.000196
28/11/2022	0.5 BX-LF-60-04		0.5	430	20	20.014	0.000196
12/09/2022	0.1 FARMACIACMT C		0.2	366	2	2.013	0.000169
17/02/2024	0.1 CMT-T02		0.2	11	2	2.02	0.0004
31/08/2021	0.5 LB-80		0.5	262	10	10.019	0.000361
28/11/2022	0.5 BX-LF-0-04		0.5	423	20	20.014	0.000196
17/11/2022	0.1 LCP-EI-114		0.5	458	37.8	37.822	0.000484
04/04/2022	0.5 SIN ID		0.5	163	15	15.018	0.000324

Figura 13. Creación de la nueva columna.

Ahora teniendo esa nueva columna se procedió a realizar una gráfica (X, Y) utilizando la columna de fecha de calibración y la nueva columna donde se realizó la resta. La gráfica obtenida se muestra en la Figura 14.

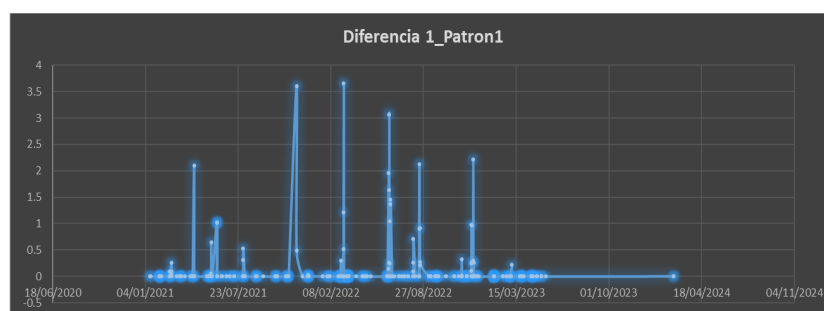


Figura 14. Gráfica (X, Y) de la columna de fecha de calibración y la columna con la diferencia.

Esto se realizó para observar los datos de manera gráfica y de acuerdo con los modelos de regresión analizados anteriormente, encontrar el modelo que tenga el gráfico similar a nuestra gráfica. Teniendo en cuenta la gráfica de árboles de regresión en la Figura 7, es el modelo que mejor se ajusta a nuestra gráfica obtenida de nuestro conjunto de datos en la Figura 14. Para comprobarlo se procedió a aplicar este algoritmo al conjunto de datos para verificar si es el algoritmo adecuado o se debe elegir otro, tomando en cuenta las métricas MSE Y MAE. Lo primero que se realizó fue crear datos de prueba y entrenamiento y crear el modelo de árbol de regresión. *Train_size=0.8*, significa que el 80% de los datos se utilizarán para entrenar el modelo y el 20% restante de prueba. Posteriormente se graficó el modelo con los datos de entrenamiento (Figura 17).

```
[ ] # Crea datos de prueba y entrenamiento
X_train, X_test, y_train, y_test = train_test_split(x, y, train_size=0.8)

[ ] # Crea el modelo de árbol de regresión
model = DecisionTreeRegressor(max_depth=5) # profundidad del árbol
model = model.fit(X_train, y_train)
```

Figura 15. Creación del modelo.

```
[10] # Graficamos para ver el modelo
plt.scatter(X_train, y_train, marker='o', c='g', alpha=0.5)
xx = np.linspace(np.min(X_train), np.max(X_train), 200).reshape(-1, 1)
yy = model.predict(xx)
plt.plot(xx, yy, '-b', alpha=0.5)
plt.title('Modelo de árbol de regresión vs datos de entrenamiento')
plt.xlabel('Punto1_Patron1')
plt.ylabel('Diferencial_Patron1')
plt.show()
```

Figura 16. Código de gráfica del modelo.

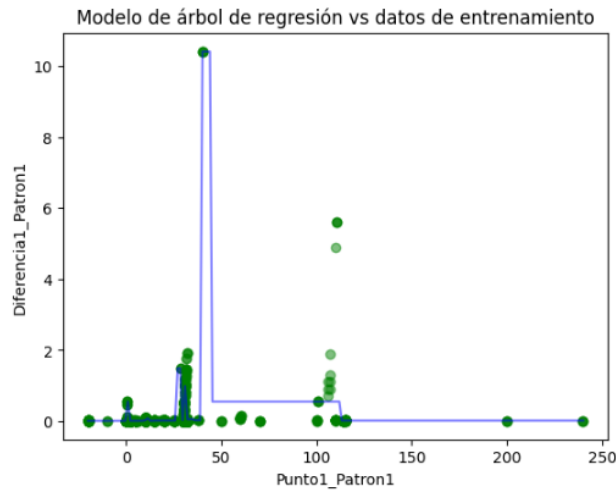


Figura 17. Gráfica de modelo vs datos de entrenamiento.

En la Figura 18, se observa que el modelo (línea azul), toca casi todos los datos de entrenamiento (puntos verdes). Y para calcular las métricas se utilizaron las funciones “mse” y “mae”.

```
[12] # Evaluación del modelo
y_pred = model.predict(X_test)
mse_score = mse(y_test, y_pred)
mae_score = mae(y_test, y_pred)
print('MSE:', '{:.2f}'.format(mse_score))
print('MAE:', '{:.2f}'.format(mae_score))
```

→ MSE: 0.12
MAE: 0.06

Figura 18. MSE Y MAE.

- MSE: 0.12
- MAE: 0.06

El Error Cuadrático Medio (MSE), es una métrica que mide el promedio de los errores cuadrados entre los valores predichos y los valores reales [14]. Unos valores cercanos a 0 indican que el modelo tiene una alta precisión y valores altos indican que el modelo tiene una baja precisión [15]. Se calcula de la siguiente manera:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

Donde:

- n es el número de observaciones.
- y_i es el valor real.
- $\hat{y}_i$ es el valor predicho.

El Error Absoluto Medio (MAE), es una métrica que mide el promedio de los errores absolutos entre los valores predichos y los valores reales [16].

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2)$$

Donde:

- n es el número de observaciones.
- y_i es el valor real.
- $\hat{y}_i$ es el valor predicho.

En los resultados obtenidos se observa que el modelo es bueno, ya que tiene buena precisión. Pero, aun así, el modelo se puede mejorar limpiando los datos, es decir se filtraron los datos atípicos (*outliers*) utilizando “*z-score*”, el cual es una medida de cuántas desviaciones estándar un punto de datos está por encima o por debajo de la media [17]. Por lo que en la Figura 19 se muestra cómo se filtraron los datos atípicos.

```
[15] # Filtrar datos atípicos (outliers) usando z-score
      z_scores = np.abs(stats.zscore(x))
      clean_mask = (z_scores < 2).all(axis=1) # Filtrar según el criterio z-score < 2
      x_clean = x[clean_mask]
      y_clean = y[clean_mask]

[16] # Generar el modelo con datos procesados
      X_train_clean, X_test_clean, y_train_clean, y_test_clean = train_test_split(x_clean, y_clean, train_size=0.8)
      model_clean = DecisionTreeRegressor(max_depth=5) # Puedes ajustar la profundidad del árbol aquí
      model_clean = model_clean.fit(X_train_clean, y_train_clean)
```

Figura 19. Filtración de datos atípicos.

Por lo que en la Figura 20 se observa la gráfica que dio como resultado el nuevo modelo.

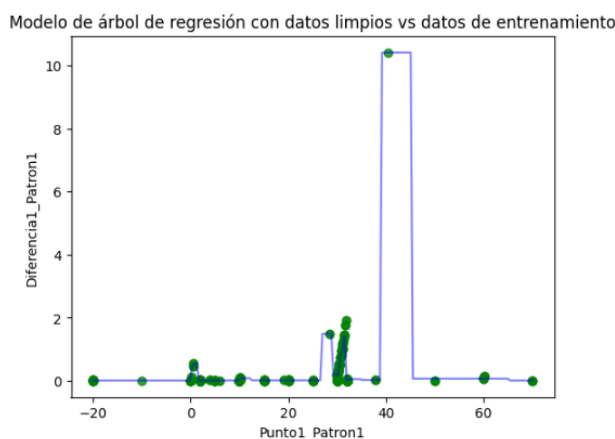


Figura 20. Modelo 2 vs datos de entrenamiento

Una vez que se limpiaron los datos y se aplicó el modelo a estos datos, se obtuvieron nuevos resultados de MSE y MAE, aplicando las mismas funciones de la Figura 21.

```
# Evaluar el modelo con datos limpios
y_pred_clean = model_clean.predict(X_test_clean)
mse_clean = mse(y_test_clean, y_pred_clean)
mae_clean = mae(y_test_clean, y_pred_clean)
print('MSE con datos limpios:', '{:.2f}'.format(mse_clean))
print('MAE con datos limpios:', '{:.2f}'.format(mae_clean))

MSE con datos limpios: 0.00
MAE con datos limpios: 0.01
```

Figura 21. Nuevos resultados de MSE y MAE.

Como se mencionó anteriormente, la precisión de las métricas MSE y MAE, es buena cuando los valores son cercanos a 0. Y al obtener un MSE con valor 0.00 y un MAE con valor de 0.01, el algoritmo es bueno.

Por lo que significa que el algoritmo de árboles de regresión es el indicado para el proyecto. Con esto se va a lograr predecir las lecturas del equipo patrón cuando el usuario vaya a realizar una calibración. La predicción va a funcionar cuando el usuario introduzca el punto que se quiere calibrar, por ejemplo 30°, y el algoritmo arrojará un resultado con varios decimales, como, por ejemplo: 0.002, esto quiere decir que al punto de calibración 30° tendrá una variación de los decimales mostrados por el modelo del algoritmo.

La predicción de la lectura del instrumento patrón va a permitir al usuario verificar si la lectura se va a encontrar en un rango aceptable con respecto al punto a calibrar, por ejemplo, si el punto a calibrar es de 30°, una

lectura válida del instrumento patrón sería 30.001° o 30.005° . Pero, si la predicción muestra valores muy grandes como por ejemplo 5, quiere decir que la lectura del instrumento patrón sería aproximadamente de 35° , y esto significa que el instrumento patrón ya no se encuentran en condiciones óptimas para calibrar instrumentos, así que el usuario debe tomar la decisión de enviar a calibrar su equipo patrón con quien corresponda y con un tiempo de anticipación aceptable con la finalidad de seguir entregando calibraciones de calidad a sus clientes.

5. Discusión

La predicción de lecturas de los instrumentos de metrología es importante ya que ofrece la oportunidad de identificar las futuras lecturas de los instrumentos y así estar preparados para una anomalía, así como también le ofrece a la empresa Metrología y Suministros de Veracruz S.A. de C.V. la oportunidad de estar preparada económicamente para calibrar en tiempo y forma el instrumento patrón. Y así seguir ofreciendo calidad y precisión en las calibraciones de los instrumentos de sus clientes.

En esta etapa se llevó a cabo la elección del algoritmo de regresión tomando en cuenta el tipo de gráfica que generaron los datos obtenidos de los documentos de Excel. Posteriormente se seleccionó el algoritmo de árboles de regresión y se evaluó mediante las métricas del Error Cuadrático Medio (MSE) y el Error Absoluto Medio (MAE).

La elección del modelo de regresión es un paso que sienta las bases para la siguiente fase del proyecto que implicará la implementación y validación del modelo seleccionado. Posteriormente se considera realizar un trabajo donde se realice el análisis comparativo de diversos algoritmos de regresión tomando en cuenta las métricas que ya se mencionaron anteriormente. En este trabajo no se realizó ese análisis comparativo ya que se pretende que en un futuro trabajo sea completamente enfocado al análisis comparativo.

Además de la evaluación del modelo de regresión seleccionado, es fundamental realizar pruebas de usabilidad del *dashboard* y la aplicación web para analizar cómo los usuarios finales interactúan con la herramienta. Estas pruebas permitirán evaluar la experiencia del usuario, identificar posibles áreas de mejora en la interfaz y medir la aceptación del usuario.

Aunque la herramienta aún se encuentra en una fase preliminar, los mockups brindan una visualización previa de la aplicación web. El patrón arquitectónico propuesto asegura que los diferentes componentes de la aplicación web se integren de manera eficiente.

6. Conclusiones

La metrología desempeña un papel fundamental en el desarrollo científico, tecnológico, económico y social, impactando sectores tan diversos como la industria, el comercio, la salud y el medio ambiente. En un contexto donde la precisión y la calidad son esenciales, la adopción de tecnologías avanzadas, como el Internet de las Cosas y la Inteligencia Artificial, se vuelve crucial, tal como lo propone la Metrología 4.0, para mejorar la eficacia en todos los procesos relacionados con la metrología.

En este artículo, se presentó un análisis detallado de los datos de calibraciones previas en la empresa Metrología y Suministros de Veracruz S.A. de C.V., utilizando métricas como MAE y MSE para seleccionar el algoritmo de regresión más adecuado. La herramienta propuesta ofrece a la empresa una ventaja competitiva al permitir la predicción de futuras lecturas en el instrumento patrón, lo que facilita la planificación proactiva y la gestión eficiente de los recursos.

El diseño del *dashboard*, junto con los *mockups* presentados, proporciona a la empresa una visualización clara y accesible de información crítica, como las lecturas históricas y proyectadas, y otros datos relevantes para la toma de decisiones. Esta visualización no solo optimiza los procesos de calibración, sino que también incrementa la confiabilidad de los resultados, apoyando a la empresa en su misión de ofrecer calibraciones de alta calidad.

Como trabajo futuro, se planea no solo desarrollar la aplicación web e integrar el algoritmo de árboles de regresión entrenado, sino también expandir la herramienta para incluir diversas funcionalidades adicionales, como la integración de otros algoritmos de *machine learning*, la personalización de los *dashboards* para diferentes usuarios, y la implementación de pruebas de usabilidad exhaustivas para asegurar que la herramienta sea intuitiva y efectiva en el entorno operativo de la empresa. Estas mejoras potenciales permitirán que la empresa no solo mantenga su posición en el mercado, sino que también avance hacia una metrología más inteligente y adaptativa.

7. Referencias

- [1] JCGM. (2020). *Guide to the expression of uncertainty in measurement-Part 6: Developing and using measurement models*. https://www.bipm.org/documents/20126/2071204/JCGM_GUM_6_2020.pdf
- [2] Cifuentes Linares, I. C. (2023). *Desarrollo de dashboard para la gestión de operaciones de la dirección de calidad de Postobón S.A.* [Tesis de Grado]. Universidad de Antioquia, Colombia. <https://hdl.handle.net/10495/33326>
- [3] De la Fuente Ruiz, S. (2019). *Implantación, estudio y control de calidad de variables climáticas en el ámbito de la metrología 4.0* [Tesis de Grado]. Universidad de Valladolid. <http://uvadoc.uva.es/handle/10324/37849>
- [4] Alberto, A., S., Jacob, F. L. (2019). *The teaching of metrology: The genesis and the epilogue of a degree*. ACM Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality, León España. <https://doi.org/10.1145/3362789.336285>
- [5] Suakanto, S., Supangkat, S. H., Suhardi, Saragih, R. (2013). *Smart city dashboard for integrating various data of sensor networks*. IEEE International Conference on ICT for Smart Society, Jakarta, Indonesia. <https://doi.org/10.1109/ICTSS.2013.6588063>
- [6] Vargas Cuervo, D. F. (2019). *Diseño e implementación de dashboards para la visualización de clústeres agregados de regiones geográficas* [Tesis de Grado]. Universidad del Bosque, Bogotá, D. C. <https://hdl.handle.net/20.500.12495/6034>
- [7] Sandoval, L. J. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica*, (11), 36-40. http://www.redicces.org.sv/jspui/bitstream/10972/3626/1/Art6_RT2018.pdf
- [8] D'angiolo, F. G., Kwist, I. F., Loiseau, M., Contreras, D. E., Asteasuain, F. (2019). *Algoritmos de Regresión Lineal aplicados al mantenimiento de un Datacenter*. XXV Congreso Argentino de Ciencias de la Computación (CACIC), Córdoba, Argentina. <https://sedici.unlp.edu.ar/handle/10915/91401>
- [9] Prashant. (2024). *Different Types of Regression Models*. <https://www.analyticsvidhya.com/blog/2022/01/different-types-of-regression-models/#>
- [10] Breiman, L., Friedman, J., Stone, C. J., Olshen, R. A. (1984). *Classification and Regression Trees*. Taylor & Francis.
- [11] Microsoft. (2024). *Patrón MVC de ASP.NET | .NET*. <https://dotnet.microsoft.com/es-es/apps/aspnet/mvc>
- [12] Trueba Espinosa, A., Camarena Sagredo, J. G., Martínez Reyes, M., López García, M. L. (2012). Automatización de la codificación del patrón modelo vista controlador (MVC) en proyectos orientados a la Web. *CIENCIA ergo-sum*, 19 (3), 239-250. <https://dialnet.unirioja.es/servlet/articulo?codigo=5199006>
- [13] Qian, Z., Kavvos, G. A., Birkedal, L. (2021). Client-server sessions in linear logic. *Proceedings of the ACM on Programming Languages*, 5 (62), 1-31. <https://doi.org/10.1145/3473567>
- [14] OA.mg. (2006). *Pattern Recognition and Machine Learning*. <https://oa.mg/work/10.1007/978-0-387-45528-0>
- [15] scikit-learn (2024). *3.4. Metrics and scoring: quantifying the quality of predictions*. https://scikit-learn.org/stable/modules/model_evaluation.html#regression-metrics
- [16] Hyndman, R. J., Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal Forecast*, 22 (4), 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- [17] SciPy. (2024). *scipy.stats.zscore — SciPy v1.13.0 Manual*. <https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.zscore.html>