



**UNIVERSIDAD NACIONAL DEL ALTIPLANO**  
**FACULTAD DE MEDICINA VETERINARIA Y**  
**ZOOTECNIA**  
**ESCUELA PROFESIONAL DE MEDICINA VETERINARIA Y**  
**ZOOTECNIA**



**ESTIMACIÓN DE LA MEDULACIÓN EN FIBRA DE**  
**ALPACA HUACAYA BLANCO, MEDIANTE**  
**ALGORITMOS DE APRENDIZAJE AUTOMÁTICO**

**TESIS**

**PRESENTADA POR:**

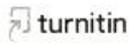
**CARLOS ALBERTO VARGAS MAMANI**

**PARA OPTAR EL TÍTULO PROFESIONAL DE:**

**MÉDICO VETERINARIO Y ZOOTECNISTA**

**PUNO – PERÚ**

**2024**



# CARLOS ALBERTO VARGAS MAMANI

## ESTIMACIÓN DE LA MEDULACIÓN EN FIBRA DE ALPACA HUACAYA BLANCO, MEDIANTE ALGORITMOS DE APRENDIZA...

Universidad Nacional del Altiplano

### Detalles del documento

Identificador de la entrega:

trn:oid::8254.416561782

98 Páginas

Fecha de entrega:

16 dic 2024, 10:48 a.m. GMT-5

14,189 Palabras

Fecha de descarga:

16 dic 2024, 10:52 a.m. GMT-5

77,099 Caracteres

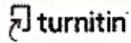
Nombre de archivo:

TESIS FORMATO CARLOS ALBERTO VARGAS MAMANI hoy lunes.docx

Tamaño de archivo:

12.4 MB





## 3% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

### Filtrado desde el informe

- Bibliografía
- Texto citado
- Texto mencionado
- Coincidencias menores (menos de 20 palabras)

### Fuentes principales

- 3% Fuentes de Internet
- 1% Publicaciones
- 1% Trabajos entregados (trabajos del estudiante)

### Marcas de integridad

#### N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

  
Domingo Ruelas Calloap  
MEDICO VETERINARIO Y ZOOTECNISTA  
C.M.V.P. 2021  
MAGISTER EN SALUD ANIMAL  
DOCTOR EN CIENCIAS DE LA SALUD

  
MSc. José Juan Quinones García  
DOCENTE FMVZ  
UNA - PUNO





## DEDICATORIA

Dedico este trabajo de investigación a la Gran Nación Aymara con territorio, idioma y cultura. Por la valentía de su población de persistir y resistir el embate de la occidentalización y colonización en el tiempo y espacio que estamos destinados a ser un estado libre y soberano.

Taqi kunans sarnaqawinakasax sarawin sarawiniwa.

**Carlos Alberto Vargas Mamani**



## AGRADECIMIENTOS

Quiero expresar mi más sincero agradecimiento a mis familiares por ser el cimiento de mi formación profesional en especial a mis progenitores.

Quiero agradecer a la Universidad Nacional del Altiplano de Puno y en particular a la Facultad de Medicina Veterinaria y Zootecnia por acogerme en sus aulas como centros del saber e investigación y contribuir en mi formación profesional, A todos los docentes de la Escuela Profesional de Medicina Veterinaria y Zootecnia por transmitir sus conocimientos y experiencias.

A las Comunidades Campesinas de Quelcaya, Chacaconiza y Chimboya por la colaboración en el desarrollo de la presente investigación como fuente de conocimientos ancestrales en la crianza de la alpaca.

A mis jurados: Dr. Clemente Vilca Castro, M.Sc. Oscar David Oros Butrón y M.Sc. Yan Pierr Manrique Quispe. Por compartirme sus conocimientos como jurados y como docentes durante mi formación profesional.

A mi asesor de tesis M.Sc. José Iván Quiñones García por su gran profesionalismo, haberme guiado y supervisado en la ejecución del presente estudio de investigación.

**Carlos Alberto Vargas Mamani**



# ÍNDICE GENERAL

	<b>Pág.</b>
<b>DEDICATORIA</b>	
<b>AGRADECIMIENTOS</b>	
<b>ÍNDICE GENERAL</b>	
<b>ÍNDICE DE TABLAS</b>	
<b>ÍNDICE DE FIGURAS</b>	
<b>INDICE DE ANEXOS</b>	
<b>ACRÓNIMOS</b>	
<b>RESUMEN .....</b>	<b>15</b>
<b>ABSTRACT.....</b>	<b>16</b>
<b>CAPÍTULO I</b>	
<b>INTRODUCCIÓN</b>	
<b>1.1. OBJETIVOS DE LA INVESTIGACIÓN.....</b>	<b>19</b>
1.1.1. Objetivo general .....	19
1.1.2. Objetivos específicos .....	19
<b>CAPÍTULO II</b>	
<b>REVISIÓN DE LITERATURA</b>	
<b>2.1. MARCO TEÓRICO .....</b>	<b>20</b>
2.1.1. Medulación.....	20
2.1.2. Tipos de medulación .....	27
2.1.3. Factores de la medulación .....	28
2.1.3.1. Factores Genéticos .....	28
2.1.3.2. Factores Morfológicos y de color .....	29
2.1.3.3. Factores Ambientales y de Manejo.....	29



2.1.4.	Algoritmos de Aprendizaje Automático .....	29
2.1.5.	Tipos de algoritmos de aprendizaje automático .....	30
2.1.5.1.	Aprendizaje supervisado .....	30
2.1.5.2.	Aprendizaje no supervisado .....	32
2.1.5.3.	Aprendizaje por refuerzo .....	32
2.1.6.	Modelos de Aprendizaje Automático para Predicciones .....	33
2.1.7.	Algoritmos usados para predicción .....	33
2.1.7.1.	AdaBoost.....	33
2.1.7.2.	Gradient Boosting .....	34
2.1.7.3.	Random Forest .....	34
2.1.7.4.	Tree .....	35
2.1.7.5.	kNN.....	35
2.1.7.6.	Neural Network.....	36
2.1.7.7.	Linear Regression .....	36

### **CAPÍTULO III**

#### **MATERIALES Y MÉTODOS**

<b>3.1.</b>	<b>LUGAR DE ESTUDIO.....</b>	<b>38</b>
<b>3.2.</b>	<b>MATERIALES DE ESTUDIO .....</b>	<b>39</b>
<b>3.3.</b>	<b>SELECCIÓN DE PREDICTORES Y MÉTRICAS DE RENDIMIENTO .</b>	<b>40</b>
<b>3.4.</b>	<b>TAMAÑO DE MUESTRA .....</b>	<b>41</b>
<b>3.5.</b>	<b>METODOLOGÍA .....</b>	<b>41</b>
3.5.1.	Selección de Alpacas.....	41
3.5.2.	Toma de Muestras de fibra.....	41
3.5.3.	Lavado de fibra .....	41
3.5.4.	Lectura de fibra en OFDA2000.....	42



3.5.5. Lectura de fibra en FiberMed.....	42
3.5.6. Generación de una Base de datos única: .....	43
3.5.7. Entrenamiento y Evaluación de Modelos de Aprendizaje Automático: .	44
<b>3.6. ANÁLISIS ESTADÍSTICO .....</b>	<b>46</b>
3.1.1. Coeficiente de Determinación $R^2$ .....	46
3.1.2. MAE (Error Absoluto Medio).....	47
3.1.3. RMSE (Raíz del Error Cuadrático Medio).....	48
<b>CAPÍTULO IV</b>	
<b>RESULTADOS Y DISCUSIÓN</b>	
<b>4.1. DIFERENTES MODELOS DE APRENDIZAJE AUTOMÁTICO PARA LA ESTIMACIÓN DE LA MEDULACIÓN TOTAL EN FIBRA DE ALPACA HUACAYA BLANCO .....</b>	<b>50</b>
<b>4.2. RENDIMIENTO DE LOS ALGORITMOS DE APRENDIZAJE AUTOMÁTICO, EN LA ESTIMACIÓN DE LA TASA DE MEDULACIÓN TOTAL EN LA FIBRA DE ALPACA HUACAYA BLANCO. ....</b>	<b>57</b>
<b>4.3. EFICACIA DEL MODELO UTILIZANDO UN CONJUNTO DE DATOS INDEPENDIENTES MEDIANTE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA LA ESTIMACIÓN DE LA MEDULACIÓN EN FIBRA DE ALPACA HUACAYA BLANCO. ....</b>	<b>60</b>
<b>4.4. DISCUSIÓN .....</b>	<b>67</b>
<b>V. CONCLUSIONES.....</b>	<b>71</b>
<b>VI. RECOMENDACIONES.....</b>	<b>72</b>
<b>VII. REFERENCIAS BIBLIOGRÁFICAS .....</b>	<b>73</b>
<b>ANEXOS.....</b>	<b>81</b>



**Área** : Producción de Camélidos Sudamericanos

**Tema** : Evaluación de Fibra de Alpaca

**Fecha de sustentación:** 19 de diciembre del 2024.



## ÍNDICE DE TABLAS

	<b>Pág.</b>
<b>Tabla 1</b> Estadísticas Descriptivas para 04 características físicas de fibra.....	50
<b>Tabla 2</b> Estadísticas Descriptivas entre Fenotipos y 04 características de fibra.....	52
<b>Tabla 3</b> Matriz de Correlaciones entre variables predictoras y la variable objetivo.	56
<b>Tabla 4</b> Comparación de modelos usando el coeficiente de determinación .....	58
<b>Tabla 5</b> Métricas de Desempeño para 7 modelos predictores de Medulación.....	61



## ÍNDICE DE FIGURAS

	<b>Pág.</b>
<b>Figura 1</b> Estructura de una fibra medulada (izquierda), frente a una fibra no medulada (derecha). .....	26
<b>Figura 2</b> Clasificación de fibra según tipo de medula. ....	28
<b>Figura 3</b> Esquematización del modelo de aprendizaje supervisado. ....	31
<b>Figura 4</b> Esquematización del modelo de aprendizaje No supervisado. ....	32
<b>Figura 5</b> Localización geográfica de las comunidades de Quelcaya, Chacaconiza y Chimboya - Distrito de Corani.....	38
<b>Figura 6</b> Flujo de trabajo desarrollado para 7 modelos predictivos de la Medulación con Orange Data Mining v. 3.37.0.....	46
<b>Figura 7</b> Gráfico de dispersión de la relación entre variables predictoras y medulación. ....	51
<b>Figura 8</b> Ploteo de la relación entre el DMF y los fenotipos.....	53
<b>Figura 9</b> Ploteo de la relación entre la Medulación y los fenotipos. ....	54
<b>Figura 10</b> Ploteo de la relación entre la DED y los fenotipos. ....	54
<b>Figura 11</b> Ploteo de la relación entre el CVD y los fenotipos. ....	55
<b>Figura 12</b> Correlograma de Spearman para variables predictoras y variable objetivo. ....	57
<b>Figura 13</b> Gráfico de barras comparativo para 3 métricas de desempeño de AA. ....	62
<b>Figura 14</b> Valores reales vs valores predichos .....	62
<b>Figura 15</b> Valores reales vs valores predichos (Random Forest).....	63
<b>Figura 16</b> Valores reales vs valores predichos linear regresión .....	64
<b>Figura 17</b> Valores reales vs valores predichos kNN .....	64
<b>Figura 18</b> Valores reales vs valores predichos AdaBoost .....	65
<b>Figura 19</b> Valores reales vs valores predichos Tree.....	66



**Figura 20** Valores reales vs valores predichos Neural Network ..... 67



## ÍNDICE DE ANEXOS

	<b>Pág.</b>
<b>ANEXO 1</b> Parametros utilizados en el software Orange Data Mining.....	81
<b>ANEXO 2</b> Panel fotográfico del uso y procesamiento de OFDA 2000 y FIBER MED. .....	93
<b>ANEXO 3</b> Declaración jurada de autenticidad de tesis .....	97
<b>ANEXO 4</b> Autorización para el deposito de tesis en el Repositorio Institucional ...	98



## ACRÓNIMOS

AA:	Aprendizaje Automático
BLUP:	Best Linear Unbiased Predictor
CFW:	Clean Fleece Weight
CVD:	Coefficiente de Variación del Diámetro
DE:	Desviación Estándar
DED:	Desviación Estándar del Diámetro
DIA:	Digital Imagen Analysis
DMF:	Diámetro Medio de Fibra
EE:	Error Estándar
FD:	Fiber Diameter
IA:	Inteligencia Artificial
MAE:	Error Absoluto Medio
MED:	Medulación
MT:	Model Tree
OFDA:	Analizador Óptico del Diámetro de Fibra.
PM:	Porcentaje de medulación
RMSE:	Raíz del Error Cuadrado Medio
SL:	Staple Length
SS:	Staple Strength
SVM:	Support Vector Machine



## RESUMEN

La presencia de médula en la fibra de alpaca, impacta negativamente en dos propiedades fundamentales: una de ellas es el confort, debido a que las fibras meduladas presentan mayor rigidez y aspereza al contacto con la piel por su incrementado diámetro y menor flexibilidad; y la segunda es el teñido, donde la cavidad medular actúa como una barrera física que impide la penetración y distribución homogénea de los colorantes, resultando en teñidos irregulares que reducen significativamente el valor comercial del producto textil final. El objetivo principal de este estudio fue evaluar el desempeño de algoritmos de aprendizaje automático supervisado para estimar la tasa de medulación total en fibra de alpacas Huacaya blanco, tui mayor. Se analizaron 439 muestras de fibra obtenidas de la región del costillar medio, de las comunidades de Quelcaya, Chacaconiza y Chimboya del distrito de Corani, Carabaya, Puno, utilizando como variables predictoras el diámetro de fibra, desviación estándar del diámetro y el coeficiente de variación del diámetro, cuyos datos fueron obtenidos del Analizador Óptico del Diámetro de Fibra - OFDA2000, mientras que la variable objetivo, (tasa de medulación total), se midió con el equipo denominado Medulador Automático - FIBERMED<sup>®</sup>. Luego del preprocesamiento de datos, se evaluaron siete algoritmos de aprendizaje automático, mediante validación cruzada en el software Orange Data Mining v. 3.37.0, utilizando el coeficiente de determinación ( $R^2$ ), la media del error absoluto (MAE) y la Raíz del Error Cuadrático Medio (RMSE), como métricas principales de desempeño. Los resultados imputaron al modelo: Random Forest ( $R^2=0.85$ , MAE=4.77 y RMSE=6.53) como el modelo más eficiente, logrando explicar el 85% de la variabilidad en la tasa de medulación total. Estos hallazgos sugieren que sí es posible estimar la medulación total en fibra de alpaca Huacaya de capa blanca utilizando herramientas de aprendizaje automático que permitiría el acceso a productores alpaqueros a un bajo costo.

**Palabras Clave:** Alpaca, Algoritmo, Aprendizaje automático, Medulación, Predictores.



## ABSTRACT

The presence of medulla in alpaca fiber negatively impacts two fundamental properties: one of them is comfort, because medulla fibers present greater rigidity and roughness in contact with the skin due to their increased diameter and lower flexibility; and the second is dyeing, where the medullary cavity acts as a physical barrier that prevents the penetration and homogeneous distribution of dyes, resulting in irregular dyeing that significantly reduces the commercial value of the final textile product. The main objective of this study was to evaluate the performance of supervised machine learning algorithms to estimate the total medulation rate in fiber from Huacaya blanco, Tui mayor alpacas. 439 fiber samples obtained from the middle rib region of the communities of Quelcaya, Chacaconiza and Chimboya in the district of Corani, Carabaya, Puno, were analyzed, using as predictor variables the fiber diameter, standard deviation of the diameter and the coefficient of variation of the diameter, whose data were obtained from the Optical Fiber Diameter Analyzer - OFDA2000, while the target variable, (total medullation rate), was measured with the equipment called Automatic Medullator - FIBERMED®. After data preprocessing, seven machine learning algorithms were evaluated, through cross-validation in the Orange Data Mining software v. 3.37.0, using the coefficient of determination ( $R^2$ ), the mean absolute error (MAE) and the Root Mean Square Error (RMSE), as main performance metrics. The results imputed the model: Random Forest ( $R^2=0.85$ ,  $MAE=4.77$  and  $RMSE=6.53$ ) as the most efficient model, managing to explain 85% of the variability in the total medullation rate. These findings suggest that it is possible to estimate the total medullation in white-coated Huacaya alpaca fiber using machine learning tools that would allow access to alpaca producers at a low cost.

**Keywords:** Alpaca, Algorithm, Machine learning, Medullation, Predictors.



# CAPÍTULO I

## INTRODUCCIÓN

La medulación en la fibra de alpaca constituye un problema crítico para la industria textil debido a que las fibras que presentan cavidades internas (totales o parciales) exhiben características desfavorables como mayor aspereza, menor flexibilidad y falta de uniformidad. Estos atributos comprometen la calidad del producto final al afectar la suavidad, el brillo, la capacidad tintórea y la eficiencia del hilado, resultando en mayores costos de producción y menor valor comercial. La reducción de la medulación es fundamental para disminuir el factor de picazón y mejorar el valor económico de las prendas (Cruz et al., 2018).

La medulación de la fibra de alpaca se ha establecido como un indicador, siendo una característica que viene posicionándose, como un indicador de la calidad y confort en la elaboración de productos textiles de la fibra de alpaca, este atributo se atribuye la presencia de la médula, siendo esto un canal celular de forma tubular ubicado que se encuentra ubicado al interior de cada fibra y que según el volumen ocupa un espacio de forma diversa, incide directamente en el confort y propiedad termorreguladora de la prenda producida. Aunque la medulación confiere propiedades únicas a la fibra de alpaca, su incidencia varía, por lo tanto, su estudio y evaluación se ha vuelto crucial para garantizar el concepto de calidad. Existen 5 tipos de medulación en la fibra de alpaca, los cuales son: La fibra fuertemente medulada, medulada continua, medulada discontinua, medulada fragmentada y finalmente la fibra no medulada (Valdivia et al., 2024).

El aprendizaje automático es una rama de la Inteligencia Artificial que se dedica a crear algoritmos que tienen la capacidad de aprender sin necesidad de programarlos o previamente programados de forma explícita. El programador alimenta datos recopilados



del campo, los selecciona, los categoriza afín de estandarizar los datos. Lo único que se debe hacer es proporcionar al algoritmo una gran cantidad de datos para que este aprenda y comprenda qué hacer en cada uno de estos escenarios. (Sandoval, 2018).

El aprendizaje automático emerge como una herramienta fundamental para predecir la medulación en fibra de alpaca, superando las limitaciones de los métodos tradicionales al procesar eficientemente múltiples variables y detectar patrones complejos. Esta tecnología, complementada con dispositivos como el FIBERMED, permite una evaluación rápida y económica de la medulación, facilitando decisiones más precisas en la selección genética y el manejo del rebaño para mejorar la calidad de la fibra (Quispe et al., 2022).

En resumen, la medulación en la fibra de alpaca representa un desafío crítico para la industria textil, afectando significativamente su calidad y valor comercial. Para abordar esta problemática, el presente estudio se centra en el desarrollo de un modelo predictivo basado en aprendizaje automático capaz de estimar con alta precisión la medulación en fibra de alpaca. Este objetivo general se aborda mediante tres objetivos específicos estratégicamente definidos: primero, el entrenamiento de diversos modelos de aprendizaje automático parámetros para la estimación de la medulación total en fibra de alpaca Huacaya blanca; segundo, la realización de un análisis comparativo exhaustivo del rendimiento de estos algoritmos para identificar el más eficiente en la predicción de la tasa de medulación; y tercero, la validación rigurosa de la eficacia del modelo seleccionado utilizando un conjunto de datos independientes. Esta aproximación metodológica busca proporcionar una herramienta confiable y accesible que permita a productores y procesadores textiles evaluar y predecir la medulación de manera precisa, contribuyendo así a la mejora de la calidad de la fibra y la optimización de los procesos productivos.



## **1.1. OBJETIVOS DE LA INVESTIGACIÓN**

### **1.1.1. Objetivo general**

Desarrollar un modelo predictivo basado en aprendizaje automático que sea capaz de estimar la medulación de la fibra de alpaca con una alta precisión.

### **1.1.2. Objetivos específicos**

- Entrenar diferentes modelos de aprendizaje automático para la estimación de la medulación total en fibra de alpaca Huacaya blanco.
- Comparar el rendimiento de los algoritmos de aprendizaje automático, en la estimación de la tasa de medulación total en la fibra de alpaca Huacaya blanco.
- Validar la eficacia del modelo utilizando un conjunto de datos independientes mediante algoritmos de aprendizaje automático para la estimación de la medulación en la fibra de alpaca Huacaya blanco.



## CAPÍTULO II

### REVISIÓN DE LITERATURA

#### 2.1. MARCO TEÓRICO

##### 2.1.1. Medulación

Si bien es cierto, el diámetro de la fibra es una de las principales características que define la calidad de la lana y fibra en los distintos procesos de producción y procesamiento textil; la presencia de fibras meduladas, también son importantes, debido a que determinan su uso en la industria e influyen directamente en la calidad de la fibra, hilados y tejidos, afectando a su vez su apariencia y rendimiento (Gupta et al., 1981; McGregor, 2006).

Las fibras meduladas se caracterizan por presentar un conducto central (médula) que contiene residuos celulares y espacios aéreos que se extienden de manera continua o fragmentada a lo largo de la estructura fibrilar (Gupta et al., 1981; Hunter, 2013; Lupton y Pfeiffer, 1998; McGregor et al., 2013). La preocupación principal se centra en las fibras con alta medulación (conocidas como fibras objetables o tipo kemp), que se distinguen por tener una médula cuyo diámetro representa más del 60% del diámetro total de la fibra.

Las fibras meduladas, principalmente las fuertemente meduladas (objetables), constituyen un defecto desde el punto de vista textil, debido al efecto sobre las propiedades del teñido en los tejidos, dado que las bolsas de aire presentes en la médula les dan una apariencia más clara que las fibras normales, afectando las propiedades ópticas de la luz que pasa a través de la fibra por



difracción y no por la diferencia de consumo de tinte por la queratina de la fibra (Aylan-Parker y McGregor, 2002).

Las fibras meduladas exhiben menor resistencia mecánica pero mayor flexibilidad, y típicamente se correlacionan con un incremento en el diámetro y una reducción en la formación de rizos (Gupta et al., 1981; Moore, 2015). Estas características impactan negativamente en las propiedades táctiles del material textil, afectando la suavidad, pilosidad, rigidez y confort al contacto con la piel en los productos finales como hilos y tejidos. Por estas razones, existe un consenso en la industria textil que considera a las fibras meduladas como elementos indeseables o contaminantes en los vellones de diversas fibras de origen animal, incluyendo lana, mohair, cachemira, alpaca y llama (Hunter, 1993; Balasingam, 2005; Frank et al., 2009; Cruz et al., 2019).

Desde varios años atrás se han realizado ingentes esfuerzos para mejorar la finura de la fibra y la lana. En el caso de Perú, las alpacas llegan a tener medias de diámetro de fibra entre 20 y 23  $\mu\text{m}$  (Montes et al., 2008; Quispe et al., 2009; Vásquez et al., 2015, Pinares et al., 2018; Cruz et al., 2019), que, a la clasificación se encuentran dentro de la calidad Alpaca Baby. Si bien, los programas de mejoramiento genético no consideran la tasa de medulación de las fibras, al tener una heredabilidad moderada, posibilita su uso como criterio de selección, lo que permitiría disminuir el efecto del factor de picazón ocasionado principalmente por fibras fuertemente meduladas (Pinares et al., 2018; Cruz et al., 2019), pero también se lograría afinar aún más la fibra. Bajo ese contexto, McGregor (2006) y Onal, et al. (2007) refieren que el grado de asociación entre la tasa de medulación y el diámetro de la fibra, en alpacas, llamas, entre otras especies productoras de lana y pelo fino. es positiva con valores que se encuentran entre moderado a alto.



Es necesario también señalar que, la medulación no es del todo censurable, dado que, la presencia de médula hace que las fibras sean más livianas, razón por la cual, tejidos en base a fibras de conejo y alpaca son más ligeras que los de lana (Czaplicki, 2012) y brindan un aislamiento térmico superior (Moore, 2015), debido a la presencia de bolsas de aire que repele el intercambio térmico (Czaplicki, 2012). Bajo esa premisa, las fibras objetables son la causa del problema de picor o picazón de fibras de alpacas, resultando de importancia reducir o eliminar este tipo de fibras, además de las que tienen medulación continua (Pinares et al., 2018).

Respecto a otras especies productoras de lana y pelo fino, la presencia de medulación en vellones de ovinos de la raza Merino es muy baja, en vellones de mohair y cachemira es de baja a moderada respectivamente, en tanto que, en fibras del buey almizclero es moderada (McGregor, 2012), en llamas, la medulación es frecuentemente alta (Martínez et al., 1997; Rodríguez et al., 2003; Poma, 2018), en vellones de alpacas, la incidencia de medulación, varía de regular a alta (McGregor, 2006; Lupton et al., 2006; Contreras, 2010; Pinares et al., 2018, 2019; Radzik et al., 2018; Cruz et al., 2019), en tanto que, en pelos de conejo Angora es considerablemente alta (McGregor, 2012) pese a contar con diámetros de oscilan entre 10 a 15 micras (Chattopadhyay et al., 2005), mostrando que, la incidencia de medulación tiende a aumentar con la edad (Martínez et al., 1997; Lupton et al., 2006, Pinares et al., 2019)

Varios estudios, respecto a la presencia de médulas, se han realizado detalladamente en lanas y fibras de mohair, utilizando el OFDA 100 y el microscopio de proyección (Balasingam, 2005); no obstante, en vellones de otras especies, existe restringida información al respecto, ya que su observación,



asistido por microscopio de proyección resulta poco confortable en tiempo y costes) y necesita mucho entrenamiento, dado que los resultados varían entre operadores (Balasingam, 2005; Shakyawar et al., 2013), sin embargo el uso del fibrómetro OFDA 100, permite obtener resultados dentro de un rango bastante amplio (Pinares et al., 2019) y en algunos casos atípicos, los parámetros de opacidad considera que deben ser redefinidos para su uso en otros tipos de fibra. Sobre el aspecto anterior, Rafat et al., (2007) y Pinares et al., (2018) exponen resultados discordantes de medulación medidos con el microscopio de proyección y el OFDA 100. Por tanto, los reportes sobre tasa de medulación en alpacas y llamas, realizados por McGregor (2006), Lupton et al., (2006), Ferguson et al., (2012) y Cruz et al., (2019), entre otros, deberían ser discutidos.

Investigaciones en fibras de alpaca, determinaron que ésta contiene fibras objetables, repercutiendo en la calidad de los productos textiles; no obstante, la fibra de alpaca es considerada de alto valor textil debido a que es suave, ligera y fina. Esto debido a que, la presencia de médulas afectaría el confort de la prenda y su uniformidad en el teñido, por tanto, las fibras meduladas se consideran objetables. Sin embargo, la presencia de fibras meduladas tienen un efecto beneficioso, pues ayudaría a la alpaca a regular su temperatura en climas fríos porque la médula retiene calor (Wang et al., 2005; Moore et al., 2011).

El porcentaje de fibras meduladas en los vellones de alpacas ha mostrado una gran variabilidad fenotípica. Esta variabilidad es mayor que la que se observa en la lana de ovino Corriedale y Romney Marsh (Villarroel, 1963; Sienna et al., 2011; Scobie et al., 2015). En fibras de 22 y 40  $\mu\text{m}$ , el porcentaje de medulación aumenta de 10 a 60%. Esta relación se puede diferenciar, toda vez que, se analice el tipo de médula de la fibra. Por lo tanto, la presencia de fibras con médula



continua y fibras fuertemente meduladas tendría un impacto en la variabilidad del diámetro de la fibra en el vellón (Villarroel, 1963; McGregor et al., 2013).

Se propuso como alternativa de solución a la presencia de médula, el descordado, para eliminar las fibras gruesas (Wang et al., 2008; Laime et al., 2016). Sin embargo, el bajo rendimiento del descordado podría explicarse por la elevada varianza del porcentaje de fibras meduladas representadas en fibras con medulación fragmentada, medulación discontinua, medulación continua y las fuertemente meduladas, así como por el hecho de que estas fibras se presentan en un amplio margen de diámetros. Sin embargo, un método alternativo que podría reducir el porcentaje de fibras meduladas es la selección (Pinares et al., 2018).

Las fibras meduladas, por su mayor resistencia a la flexión y presión, constituyen un problema para la industria textil, por lo que es fundamental contar con instrumentos prácticos para identificarlas (Quispe, M.D. 2023).

Las células de la médula pueden romperse por completo durante el proceso de queratinización, lo que puede dejar un canal hueco en el centro de la fibra (Balasingam, 2005).

Las características de la fibra de alpaca desde un enfoque textil, aumentarían su calidad, si se logra reducir o eliminar el factor de picazón, aumentando su precio en el mercado. La reducción del porcentaje de medulación por selección, que podría combinarse con procesos de depilación mecánica, puede reducir el picor producido principalmente por las fibras fuertemente meduladas. Esto mejoraría la posición de la fibra de alpaca en el mercado textil de las fibras naturales, mejorando el consumo respecto a otras fibras nobles. Esto aumentaría la motivación y el impulso de toda la cadena productiva, con un impacto



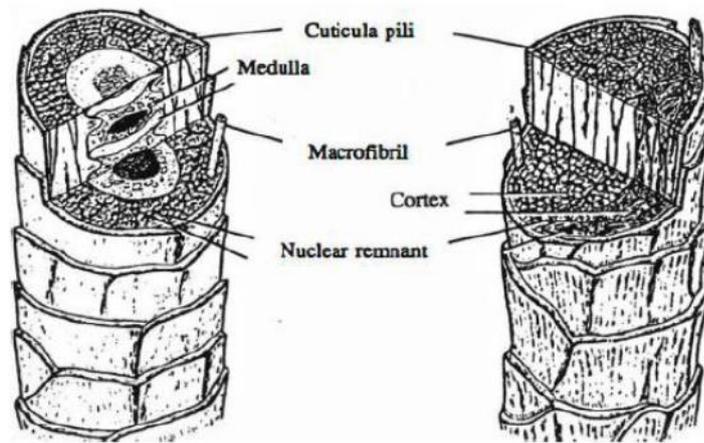
económico en los productores de alpaca menos favorecidos. Respecto a la estructura morfológica de la fibra (Fig.1) podemos señalar que, la parte central de la fibra es la médula, que tiene un vacío medular (aire) continuo o fragmentado a lo largo de la corteza (Villaroel, 1963; Carpio, 1991 y Czaplicki, 2012).

La capa ubicada en la zona intermedia que proporciona estructura a la fibra (la corteza), aumenta su proporción relativa a medida que disminuye su diámetro. Las fibras más finas solo tienen la corteza y la cutícula; sin embargo, las células corticales representan más del 90% de la fibra, lo que es comparable a las lanas finas de los ovinos (Villaroel, 1963 y Carpio, 1991). Las células corticales (células blandas y duras, respectivamente) se distribuyen bilateralmente en las fibras finas. Este hecho parece estar estrechamente relacionado con la presencia de rizos en la fibra de alpaca Huacaya. Sin embargo, en la alpaca Suri, no hay una distinción cortical clara entre la ortocorteza y la paracorteza, por lo que esto no ocurre en ella (Villaroel, 1963 y Carpio, 1991).

Las células de la cutícula están superpuestas, como las escamas de un pez, para formar la cutícula. La capa exterior de toda la fibra se conoce como cutícula. La epicutícula, la exocutícula y la endocutícula son las tres capas que componen cada celda de la cutícula. La alpaca Huacaya tiene una forma áspera y tiene un número diferente de cutículas que las de la alpaca Suri (Villaroel, 1963; Carpio, 1991 y Valbonesi et al., 2010).

### Figura 1

*Estructura de una fibra medulada (izquierda), frente a una fibra no medulada (derecha).*



Debido a que el aire se atrapa dentro de la médula y entre las fibras (mínima conductividad térmica), el vellón de alpacas puede beneficiarse de la medulación como aislante natural haciendo función termorreguladora (Moore et al., 2011). Esto resulta en un efecto aislante (aislamiento térmico inverso de la conductancia), que es la medida de la capacidad para reducir el flujo de calor. Por lo tanto, la fibra de alpacas se considera un mejor material aislante que la lana debido a su alto contenido de vacío medular (Wang et al., 2005).

La medulación es un problema importante para la industria textil durante su industrialización, especialmente en el teñido, ya que la médula hace que las fibras meduladas parezcan más claras que las fibras no meduladas debido a su mayor refracción de luz. Esto se debe a que la médula influye en las propiedades ópticas de la luz que pasa a través de la fibra por difracción, no a las diferencias en la absorción de colorante por la queratina de la fibra (Hunter, 1993). Sin embargo, según Quispe en el 2014, en su investigación, señala que las fibras



meduladas o parcialmente meduladas no absorben totalmente los colorantes (Quispe, 2014).

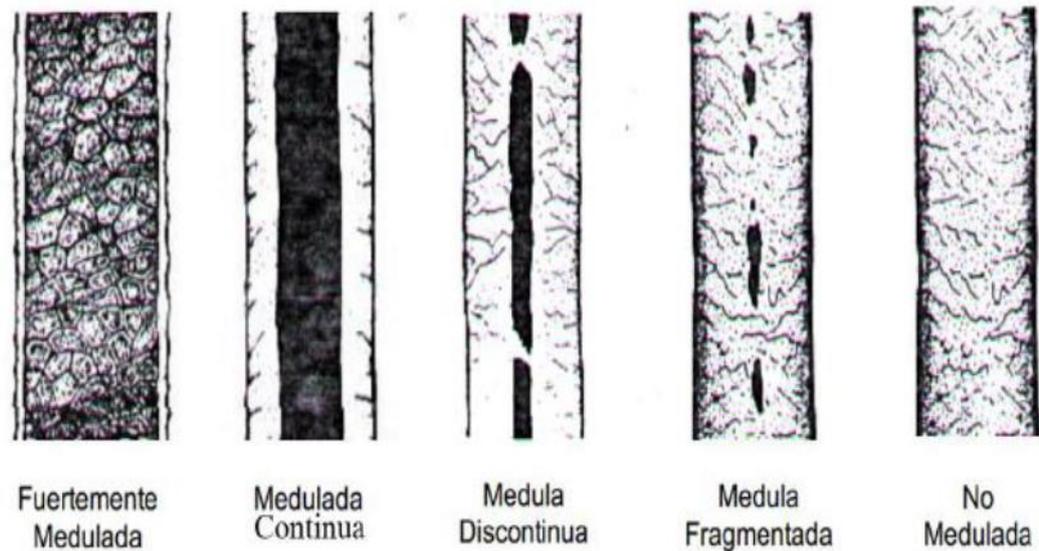
La industria considera que las fibras meduladas son un contaminante en el procesamiento textil. Las bragas de vellón contienen más del 33% de fibras meduladas (Hack et al., 1999), y la medulación varía según la región del vellón (Aylan-Parker y McGregor 2002). Por último, pero no menos importante, la medulación puede tener un impacto negativo significativo en el valor textil y su uso final, especialmente en el hilado. No obstante, la medulación no siempre se considera un defecto, ya que la presencia de médula en las fibras hace que sean más livianas, lo que explica por qué los tejidos hechos con fibras de conejo y alpaca son más ligeros que los hechos con lana (Czaplicki, 2012)

### **2.1.2. Tipos de medulación**

De acuerdo con el plano de observación longitudinal, la médula de la fibra de alpaca presenta una variedad de características: a) no medulada; b) fragmentada, en la que la médula solo se encuentra fragmentada de vez en cuando en intervalos irregulares; c) discontinua, en la que la médula se muestra en intervalos bastante regulares a lo largo de la longitud de la fibra; d) continua, en la que la médula se muestra a lo largo de la longitud de la fibra; e) fuertemente medulada, en la que las células de la parte central de la fibra se compactan para dar una apariencia tubular (Villarroel, 1963).

## Figura 2

*Clasificación de fibra según tipo de medula.*



Fuente: Villaroel, (1963).

### 2.1.3. Factores de la medulación

La presencia de medulación en la fibra de alpaca y lana de ovino está influenciada por varios factores, entre ellos destacan los factores genéticos, morfológicos y ambientales.

#### 2.1.3.1. Factores Genéticos

La medulación en la fibra de alpaca tiene una heredabilidad de moderada a alta, lo que sugiere que es posible seleccionar genéticamente para reducir la medulación. En alpacas Huacaya, se estimó la heredabilidad de la medulación en 0.225 y en Suri en 0.664 (Cruz et al., 2018). Por otro lado, la correlación genética entre la medulación y el diámetro de la fibra es alta, lo que indica que la selección para reducir la medulación también puede reducir el diámetro de la fibra (Pinares et al., 2018).



### **2.1.3.2. Factores Morfológicos y de color**

En alpacas, los machos tienden a tener fibras más gruesas y un mayor grado de medulación en comparación con las hembras. Además, la fibra de color oscuro presenta una mayor medulación y fibras más gruesas en comparación con la fibra de color claro (Radzik-Rant & Wiercińska, 2021). La medulación también varía según la ubicación en el cuerpo del animal. Las áreas como el cuello y las patas tienden a tener una mayor medulación y fibras más gruesas en comparación con la zona media del cuerpo (Radzik-Rant, et al., 2021).

### **2.1.3.3. Factores Ambientales y de Manejo**

La incidencia de la medulación también varía con la edad del animal y la localización geográfica. Sin embargo, una vez que se considera el diámetro promedio de la fibra, ni la edad ni el sexo resultan ser factores significativos, en tanto que, las diferencias en la medulación entre diferentes ubicaciones pueden explicarse por programas de selección genética específicos para reducir el diámetro de la fibra (Quispe et al., 2022).

### **2.1.4. Algoritmos de Aprendizaje Automático**

Un algoritmo es un conjunto de reglas o instrucciones definidas y ordenadas que permiten realizar una tarea o resolver un problema (Cormen et al., 2009). Por otro lado, el aprendizaje automático es un campo de la inteligencia artificial que se enfoca en desarrollar algoritmos capaces de aprender y hacer predicciones por sí mismos. A diferencia de la programación tradicional, donde el desarrollador debe codificar manualmente todas las reglas y excepciones, en el aprendizaje automático sólo se necesita proporcionar un gran volumen de datos al



algoritmo para que éste pueda detectar patrones y aprender a tomar decisiones basadas en esos datos. Dentro del aprendizaje automático, existen dos enfoques principales: el aprendizaje supervisado, donde los datos de entrenamiento incluyen las respuestas deseadas, y el aprendizaje no supervisado, donde el algoritmo debe encontrar patrones y estructuras por sí mismo sin respuestas predefinidas (Sandoval, 2018).

### **2.1.5. Tipos de algoritmos de aprendizaje automático**

Existen 3 modelos de aprendizaje automático, estos son: Aprendizaje Supervisado, Aprendizaje No Supervisado y Aprendizaje por Refuerzo, que se detalla a continuación.

#### **2.1.5.1. Aprendizaje supervisado**

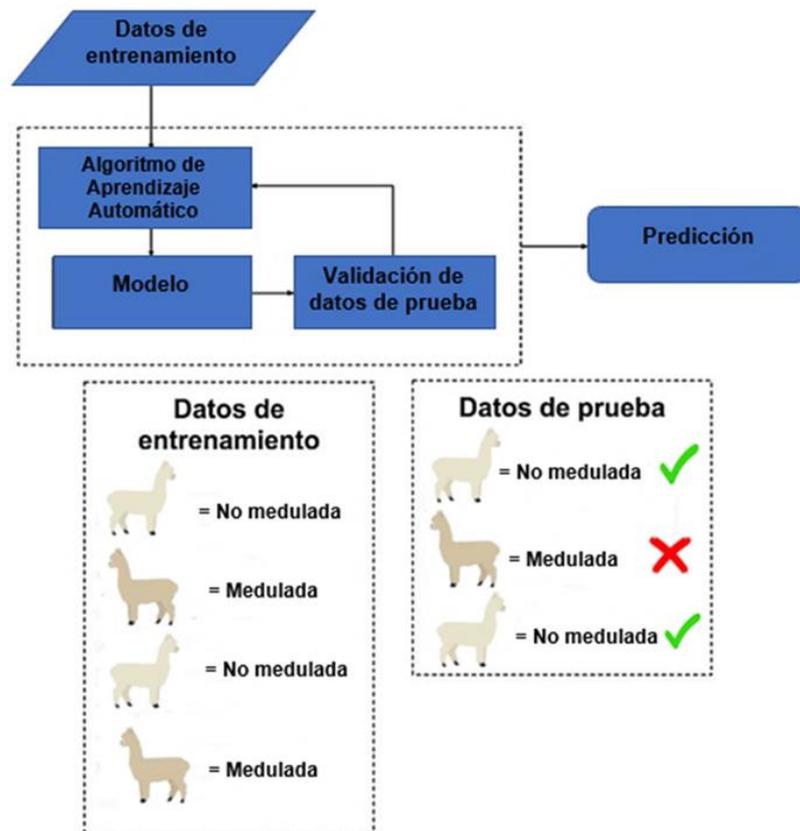
En el aprendizaje automático supervisado, se entrena un algoritmo proporcionándole un conjunto de datos que incluye tanto las características o variables de entrada como las respuestas u objetivos deseados (etiquetas). De esta manera, el algoritmo aprende a mapear las características con las etiquetas correspondientes. Una vez entrenado, el algoritmo puede hacer predicciones sobre nuevos datos, utilizando únicamente las características para inferir las respuestas. Existen dos tipos principales de algoritmos supervisados: los de clasificación, que predicen etiquetas categóricas, y los de regresión, que predicen valores numéricos continuos. En este enfoque, el algoritmo debe descubrir por sí mismo patrones y estructuras inherentes en los datos, agrupándolos según sus similitudes o siguiendo ciertos parámetros definidos por el operador. El

objetivo es encontrar relaciones y categorías ocultas dentro del conjunto de datos, sin conocimiento previo de las etiquetas. (Sandoval, 2018).

Cada registro está marcado con un valor de la variable objetivo, y se utilizan diversas técnicas que permiten anticipar el valor de dicha variable en un nuevo registro. Una vez entrenado el algoritmo, debe tener la habilidad de anticipar el valor de esa variable para un episodio que no se haya mostrado al sistema anteriormente. Así los datos de entrenamiento y los de prueba deben ser distintos y ser guardados de manera individual. Se aplica en trabajos de categorización, pronósticos e identificación de similitudes. (Núñez et al., 2019).

**Figura 3**

*Esquematación del modelo de aprendizaje supervisado.*



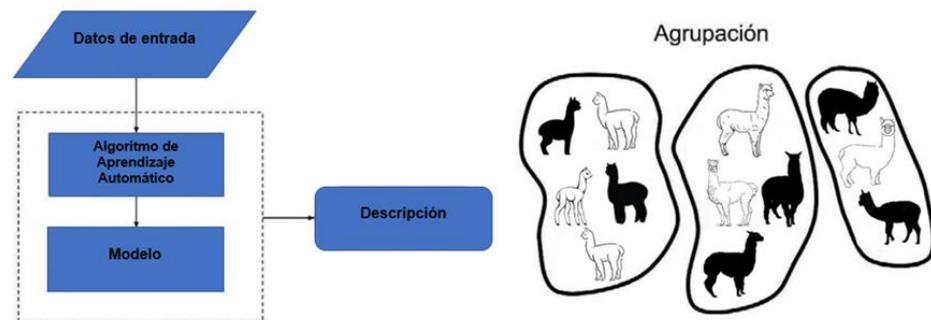
### 2.1.5.2. Aprendizaje no supervisado

Por otro lado, el aprendizaje automático no supervisado solo requiere proporcionar las características al algoritmo, sin incluir las etiquetas o respuestas deseadas (Cormen et al., 2009).

Consiste en identificar patrones o tendencias en los datos sin utilizar una variable objetivo. Es muy empleado para clasificar muestras en grupos de manera automática y para reducir el número de variables y la complejidad de los modelos (Núñez et al., 2019).

#### Figura 4

*Esquemmatización del modelo de aprendizaje No supervisado.*



### 2.1.5.3. Aprendizaje por refuerzo

En este tipo de aprendizaje, se ingresa al sistema en un ambiente dinámico, se proporcionan datos etiquetados, así como sin etiquetar, el sistema interactúa con el entorno y recibe recompensas negativas o positivas de acuerdo con sus acciones permitiendo que perfeccione el modelo, mejorando la caracterización y/o las clasificaciones (Vega et al., 2020).

### 2.1.6. Modelos de Aprendizaje Automático para Predicciones

Los modelos más utilizados para realizar predicciones en aprendizaje automático incluyen la **Regresión Lineal**, que ajusta una línea para predecir valores continuos; **k-Nearest Neighbors (kNN)**, que predice según los vecinos más cercanos; **Árboles de Decisión**, que dividen los datos en nodos basados en características; **Random Forest**, un ensamble de árboles que mejora la precisión; y **Gradient Boosting**, que construye árboles secuenciales para corregir errores. Otros modelos clave son **Support Vector Machines (SVM)**, que busca un hiperplano óptimo para separar clases; **Redes Neuronales** para tareas más complejas; **AdaBoost**, que refuerza clasificadores débiles; **Naive Bayes**, ideal para clasificación de texto; y la **Regresión Logística** para clasificación binaria. Cada modelo tiene sus ventajas dependiendo de la tarea, los datos y los recursos disponibles (Pertuz, 2022).

### 2.1.7. Algoritmos usados para predicción

A continuación, se presenta un resumen de los principales algoritmos utilizados en el presente trabajo de investigación.

#### 2.1.7.1. AdaBoost

(Adaptive Boosting) funciona combinando múltiples clasificadores débiles (como árboles de decisión simples) para crear un modelo fuerte. En cada iteración, AdaBoost ajusta el peso de las muestras incorrectamente clasificadas, enfocándose más en aquellas que los clasificadores anteriores no pudieron predecir correctamente. Al final, el modelo realiza predicciones ponderando las decisiones de todos los clasificadores débiles según su precisión. Los clasificadores que tuvieron mejor rendimiento en



predicciones tienen un mayor peso en la predicción final, logrando así un modelo robusto y preciso (Pertuz, 2022; Mirajlili y Raschka, 2020).

### **2.1.7.2. Gradient Boosting**

Es un modelo de ensamble que construye predictores secuencialmente, cada uno corrigiendo los errores del anterior. A diferencia de Random Forest, que utiliza árboles independientes, Gradient Boosting entrena árboles de decisión de manera secuencial. En cada paso, el modelo ajusta un nuevo árbol para minimizar el error de los árboles anteriores, utilizando gradientes (derivadas) para ajustar la predicción. El proceso comienza con un modelo inicial simple (generalmente una predicción promedio) y luego, en cada iteración, se calcula la diferencia entre las predicciones actuales y los valores reales (residuales). Un nuevo árbol es entrenado para predecir esos residuales, y las predicciones del conjunto de árboles se actualizan sumando las predicciones ponderadas de este nuevo árbol. Este proceso continúa hasta que se alcanza un número determinado de iteraciones o el error se minimiza lo suficiente (Perluz, 2022).

Así, Gradient Boosting crea un modelo robusto, capaz de realizar predicciones más precisas al enfocarse en corregir los errores acumulados en cada iteración (Pertuz, 2022; Mirajlili y Raschka, 2020).

### **2.1.7.3. Random Forest**

Es un modelo de aprendizaje automático que realiza predicciones combinando el resultado de múltiples árboles de decisión independientes. Cada árbol en el bosque es entrenado con una muestra aleatoria de los



datos (con reemplazo, un proceso conocido como "bagging") y, además, a cada división de los árboles se les permite usar solo un subconjunto aleatorio de características. Para realizar una predicción, cada árbol en el bosque genera su predicción individual, y el Random Forest toma la decisión final mediante una votación mayoritaria en clasificación o promediando las predicciones en el caso de regresión. Este enfoque mejora la precisión del modelo y reduce el riesgo de sobreajuste (Pertuz, 2022; Mirajlili y Raschka, 2020).

#### 2.1.7.4. Tree

(Árbol de decisión) predice siguiendo una estructura jerárquica donde los datos se dividen secuencialmente en subconjuntos basados en preguntas binarias sobre sus características. Comienza en la raíz del árbol, donde se evalúa una característica del dato de entrada contra un umbral, y según el resultado, se sigue por una rama hacia nodos hijos. En cada nodo interno se aplica una nueva condición hasta llegar a una hoja, donde se asigna una predicción final. Para clasificación, la hoja asigna una clase, y para regresión, devuelve un valor promedio de los datos en esa hoja (Pertuz, 2022; Mirajlili y Raschka, 2020).

#### 2.1.7.5. kNN

El modelo **k-Nearest Neighbors** (kNN) realiza predicciones basándose en la similitud entre datos. Cuando recibe un nuevo dato para predecir, kNN busca los **k** puntos de datos más cercanos en el conjunto de entrenamiento, utilizando una medida de distancia, como la distancia euclidiana. Para clasificación, el modelo asigna la clase que es más



frecuente entre los  $k$  vecinos cercanos; mientras que, en regresión, el valor de la predicción es el promedio de los valores de los vecinos más cercanos. KNN es un algoritmo simple pero efectivo, donde el rendimiento depende de la elección adecuada de  $k$  y la métrica de distancia utilizada (Pertuz, 2022; Mirajlili y Raschka, 2020).

#### **2.1.7.6. Neural Network**

Una red neuronal (perceptrón) realiza predicciones procesando los datos de entrada a través de capas interconectadas de neuronas artificiales. Cada neurona recibe entradas, las multiplica por pesos asignados, suma estos productos, y pasa el resultado por una función de activación. Este proceso se repite en cada capa, propagando la información hacia adelante. Los pesos entre las neuronas se ajustan durante el entrenamiento para minimizar el error entre las predicciones y los valores reales. En la predicción, la red utiliza estos pesos optimizados para procesar nuevos datos de entrada, transformándolos capa por capa hasta producir una salida final. La complejidad de las relaciones que la red puede modelar aumenta con el número de capas y neuronas, permitiéndole capturar patrones intrincados en los datos y realizar predicciones basadas en estos patrones aprendidos (Pertuz, 2022; Mirajlili y Raschka, 2020).

#### **2.1.7.7. Linear Regression**

El modelo de regresión lineal realiza predicciones estableciendo una relación lineal entre las variables de entrada (predictores) y la variable de salida (objetivo). Durante el entrenamiento, el modelo determina los coeficientes óptimos para cada predictor y un término de intercepto,



minimizando la suma de los errores cuadráticos entre las predicciones y los valores reales. En la fase de predicción, el modelo toma los valores de las variables de entrada para una nueva observación, multiplica cada uno por su coeficiente correspondiente, suma estos productos junto con el intercepto, y produce un valor predicho para la variable objetivo. La simplicidad de este enfoque lo hace interpretable y computacionalmente eficiente, aunque limita su capacidad para capturar relaciones no lineales complejas en los datos (Pertuz, 2022; Mirajlili y Raschka, 2020).

## CAPÍTULO III

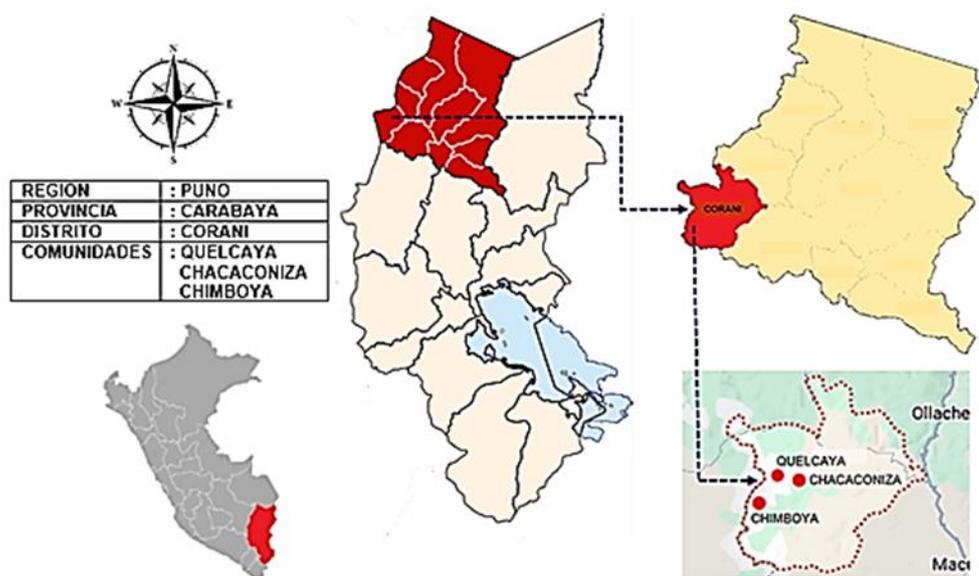
### MATERIALES Y MÉTODOS

#### 3.1. LUGAR DE ESTUDIO

La investigación se llevó a cabo en las comunidades de Quelcaya; 14°00'31" Latitud Sur, 70°44'26" Longitud Oeste, Chacaconiza; 14°01'12" Latitud Sur, 70°39'48" Longitud Oeste, Chimboya; 14°02'58" Latitud Sur, 70°45'47" Longitud Oeste, ubicados en el distrito de Corani, provincia de Carabaya, región Puno, a 311 Km. de distancia de la capital de la región, con una altitud que oscila entre 4406 m.s.n.m. hasta los 4794 m.s.n.m. una superficie de 852.99 km<sup>2</sup> con una población aproximada de 3622 habitantes (INEI, 2012).

#### Figura 5

*Localización geográfica de las comunidades de Quelcaya, Chacaconiza y Chimboya - Distrito de Corani.*





### 3.2. MATERIALES DE ESTUDIO

Para el presente estudio, se utilizaron los siguientes materiales de campo para la toma de muestras de fibra de alpaca y caracterización fenotípica:

- Tijeras
- Bolsas de plástico de 3x5cm
- Lapiceros indelebles
- Tableros
- Etiquetas
- Registros de caracterización de Fenotipo.

Respecto a materiales y equipos de laboratorio que se utilizaron para el análisis de fibra de alpaca, se usaron los siguientes:

- Alcohol isopropílico
- Bencina
- Papel Toalla
- Microtomo de Hardy
- Pipeta
- Aceite de inmersión
- Cuchillas
- Láminas porta y cubre objetos
- Analizador óptico del diámetro de fibra OFDA2000.
- Medulador Automático FIBERMED.

### 3.3. SELECCIÓN DE PREDICTORES Y MÉTRICAS DE RENDIMIENTO

Para el presente estudio, se consideró el uso de 03 predictores numéricos siendo estos el diámetro medio de la fibra (DMF), la desviación estándar del diámetro (DED) y el coeficiente de variación del diámetro (CVD) obtenidos de las lecturas en el analizador óptico del diámetro de fibra (OFDA2000) y 01 predictor categórico (Fenotipo) obtenido de un registro de caracterización; para predecir la variable objetivo (Medulación) a través del uso del mejor modelo de aprendizaje automático considerando las métricas de desempeño como el coeficiente de determinación ( $R^2$ ), la Media del Error Absoluto (MAE) y la Raíz del Error Cuadrático Medio (RMSE).

$$MED_i = \beta_0 + \beta_1 \cdot DMF_i + \beta_2 \cdot DED_i + \beta_3 \cdot CVD_i + \sum_{j=1}^{k-1} \gamma_j \cdot Fenotipo_{ij} + \epsilon_i$$

Donde:

- $MED_i$ : Medulación total de la muestra  $i$  (variable respuesta).
- $\beta_0$ : Intercepto del modelo.
- $\beta_1, \beta_2, \beta_3$ : Coeficientes de regresión asociados a las variables continuas (DMF, DED y CVD, respectivamente).
- $\gamma_j$ : Coeficientes asociados a los niveles de Fenotipo, siendo  $j$  los diferentes niveles de esta variable categórica. Si hay  $k$  niveles en la variable Fenotipo, se incluye  $k - 1$  variables indicadoras (dummy variables).
- $\epsilon_i$ : Término de error aleatorio (normalmente distribuido con media 0 y varianza constante).

Para el entrenamiento y validación se utilizó el enfoque de validación cruzada de 3 pliegues: Es importante indicar que luego del preprocesamiento de datos, se obtuvo 439 datos válidos que se dividieron en 3 grupos, donde 2 grupos (292 datos aleatorios) se utilizaron para entrenamiento de los modelos de aprendizaje automático y un grupo (147 datos aleatorios) para validación de modelos.



### **3.4. TAMAÑO DE MUESTRA**

Se emplearon 439 muestras de fibra de alpaca Huacaya de color de capa blanco “Tui mayor”, provenientes de las comunidades de Quelcaya, Chimboya y Chacaconiza. Las muestras fueron obtenidas de la región del costillar medio, en una cantidad aproximada entre 2 – 4 gramos, siendo etiquetadas y almacenadas en bolsas de plástico considerando el arete, nombre del propietario, comunidad, raza, dentición y sexo de cada animal.

### **3.5. METODOLOGÍA**

#### **3.5.1. Selección de Alpacas**

Las alpacas seleccionadas provienen de los Centros de Producción de Reproductores (CPRs) de las comunidades de Chimboya, Chacaconiza y Quelcaya, que participan en el programa de asistencia técnica del proyecto de alpacas impulsado por la Municipalidad Distrital de Corani. Dentro de este grupo, se eligieron únicamente alpacas Huacaya de pelaje blanco, categoría tui mayor de aproximadamente 2 años de edad, identificadas por la presencia de dos incisivos permanentes (Huanca, 1996).

#### **3.5.2. Toma de Muestras de fibra**

Las muestras se obtuvieron de la región del costillar medio en una cantidad de 2 - 3 gramos, luego fueron etiquetadas y rotuladas considerando datos del animal y el propietario. Las muestras se obtuvieron en el mes de septiembre, antes de la esquila.

#### **3.5.3. Lavado de fibra**

Para este procedimiento, se utilizó un recipiente de 1 litro de capacidad en el que se mezclaron alcohol isopropílico y bencina en proporciones iguales, es



decir, 500 ml de cada solvente, según las recomendaciones del fabricante del medulador automático. Las mechas de fibra obtenidas fueron lavadas empleando pinzas de sujeción, realizando movimientos laterales dentro del recipiente durante un tiempo aproximado de 1 a 2 minutos por muestra. Posteriormente, las muestras se enjuagaron con agua embotellada comercial durante 15 segundos. Una vez lavadas, las muestras fueron secadas utilizando papel toalla, para luego proceder con su lectura en el OFDA2000.

#### **3.5.4. Lectura de fibra en OFDA2000**

Una vez secas las muestras, se procedió a lecturar las mechas de fibra lavadas bajo el modo de mecha entera (Staple Mode), siguiendo las recomendaciones del fabricante, que se resume en los siguientes pasos:

- a. Encendido del equipo OFDA2000
- b. Ingreso al programa Meswin y al ejecutable OFDA2000.
- c. Verificación del estado de calibración del equipo mediante el uso de Poliéster estándar de 18.9 micras.
- d. Creación de carpeta de trabajo “Variables Predictoras”.
- e. Lectura de 500 muestras de fibra de alpaca lavada.
- f. Exportación de resultados: “Variables\_Predictoras.csv”

#### **3.5.5. Lectura de fibra en FiberMed**

Las muestras de fibra de alpaca previamente lecturadas en el equipo OFDA2000, fueron recuperadas para analizarlas en el medulador automático denominado FIBERMED siguiendo los siguientes pasos:

- a. Se leió 8 patrones de lana lavada y cardada para determinar el estado de calibración del equipo conforme a las recomendaciones del fabricante.



- b. Las mechas de fibra lecturadas en OFDA2000, fueron fragmentadas utilizando un micrótopo de Hardy para obtener “Snippets” de 10 – 20 mm.
- c. Los fragmentos fueron colocados en una lámina porta objetos luego se agregó una gota de aceite de inmersión sobre los fragmentos de fibra con el objetivo de visualizar con mejor nitidez los tipos de médula.
- d. Una vez preparada la muestra de fibra, se colocó con lapicero indeleble el ID de cada muestra para ingresarla al software del medulador automático.
- e. Para lecturar los fragmentos de fibra en el medulador automático, se procedió a encender el equipo y luego crear el archivo para el almacenamiento de los resultados con el nombre “Medulacion\_Corani\_xlsx”
- f. Una vez creado el archivo de almacenamiento, se procedió con la configuración del equipo considerando la ecuación polinómica diseñado para fibra de alpaca.
- g. Posteriormente se procedió con las lecturas de las 500 muestras de fibra de alpaca, ingresando el identificad (Nro de Arete) de cada una de ellas.
- h. Finalmente se realizó la exportación de resultados contenidos en el archivo “Medulacion\_Corani\_xlsx”.

### **3.5.6. Generación de una Base de datos única:**

En los pasos previos, se generaron dos archivos individuales correspondientes a cada equipo de medición de fibra. El primero, titulado “Variables\_Objetivo.xlsx”, generado por FIBERMED, mientras que el segundo, denominado “Variables\_Predictoras.csv”, fue producido por OFDA2000. Con el objetivo de evaluar siete modelos de aprendizaje automático, ambos archivos fueron combinados en un solo conjunto de datos. En esta fusión se incluyeron los regresores DMF, DED y el CVD. Además, se integró el Porcentaje de Medulación

Total (MED), que fue extraído como variable objetivo a partir de las lecturas realizadas con el medulador FIBERMED. El archivo final, denominado “Alpaca\_Med.xlsx”, contiene un total de 500 datos de mediciones de fibra de alpaca, con los valores correspondientes a DMF, DED, CVD, MED y Fenotipo, que posteriormente fueron utilizados para la evaluación de los algoritmos de aprendizaje automático.

### **3.5.7. Entrenamiento y Evaluación de Modelos de Aprendizaje Automático:**

Antes del entrenamiento, se llevó a cabo el preprocesamiento de los datos, que incluyó la eliminación de valores atípicos y registros con datos faltantes. De un total de 500 registros, se validaron 439. Los modelos de aprendizaje automático fueron entrenados y evaluados utilizando los parámetros predeterminados del programa **Orange Data Mining v. 3.37.0**. El conjunto de datos fue evaluado mediante validación cruzada con 3 grupos de datos, donde 2 grupos de datos se utilizaron para el entrenamiento y 1 grupo para la validación. A continuación, se resumen los parámetros empleados para los principales algoritmos de aprendizaje automático.

**Linear Regression:** La configuración de regresión lineal en Orange se realizó, con los siguientes parámetros: **Fit intercept:** Está activado, lo que significa que se incluye un intercepto en el modelo. Si se desactiva, el intercepto será cero. **Regularization:** Se selecciona "No regularization", lo que indica que no se aplican restricciones en los coeficientes.

**Neural Network:** Se utilizaron los siguientes parámetros para la **red neuronal** en Orange: **Neurons in hidden layers:** 120, lo que indica que hay 120 neuronas en la capa oculta. **Activation:** ReLU (Rectified Linear Unit), utilizada para activar las neuronas. **Solver:** Adam, un optimizador avanzado para ajustar



los pesos. **Regularization ( $\alpha = 0.0001$ )**: Controla el grado de regularización para evitar sobreajuste. **Maximal number of iterations**: 200, el número máximo de iteraciones para entrenar el modelo. **Replicable training**: Activado, asegura la repetibilidad del entrenamiento.

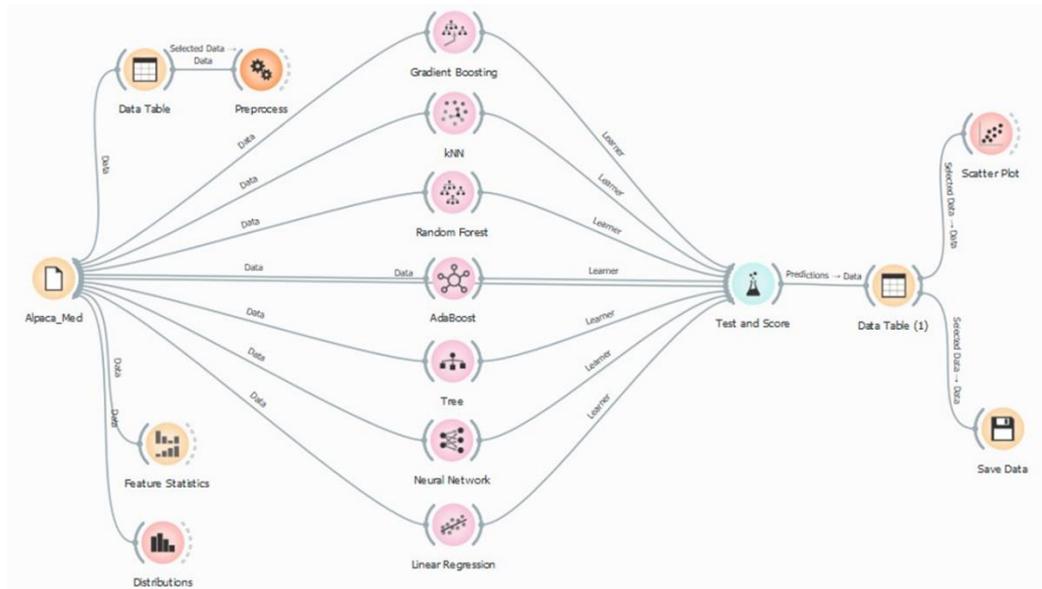
**Random Forest**: Los parámetros utilizados fueron: **Number of trees**: 50, el número de árboles en el bosque. **Number of attributes considered at each split**: 5, el número de atributos evaluados en cada división del árbol. **Limit depth of individual trees**: 3, para limitar la profundidad de los árboles. **Do not split subsets smaller than**: 5, el tamaño mínimo de los subconjuntos antes de detener la división.

**Gradiente Boosting**: Los parámetros utilizados fueron: Método: Gradient Boosting (scikit-learn), Número de árboles: 105, Tasa de aprendizaje (Learning rate): 0,100; Entrenamiento replicable: Activado, Límite de profundidad de árboles individuales: 3; No dividir subconjuntos menores que: 2; Fracción de instancias de entrenamiento: 1,00; Aplicar automáticamente: Activado.

**kNN**: Número de vecinos más cercanos: 20, Métrica: Euclidean, Peso (Weight): By Distances Indica cómo se ponderan los vecinos al hacer predicciones. "By Distances" sugiere que los vecinos más cercanos tienen más influencia.

**Figura 6**

*Flujo de trabajo desarrollado para 7 modelos predictivos de la Medulación con Orange Data Mining v. 3.37.0*



### 3.6. ANÁLISIS ESTADÍSTICO

Las métricas **R<sup>2</sup> (Coeficiente de determinación)**, **MAE (Error Absoluto Medio)** y **RMSE (Raíz del Error Cuadrático Medio)** fueron utilizadas para evaluar el rendimiento de siete modelos predictivos en el contexto del aprendizaje automático. Cada una de estas métricas proporciona información específica sobre qué tan bien se ajusta un modelo a los datos observados.

#### 3.1.1. Coeficiente de Determinación R<sup>2</sup>

El **R<sup>2</sup>** es una métrica que indica la proporción de la varianza en la variable dependiente (objetivo) que es explicada por el modelo. Se usa para evaluar la calidad del ajuste de un modelo de regresión. Su valor varía entre (0 y 1) en algunos casos, puede ser negativo si el modelo es extremadamente ineficiente (Mirajlili y Raschka, 2020). Donde:

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

- $y_i$ : Valor real de la variable dependiente.
- $\hat{y}_i$ : Valor predicho por el modelo.
- $\bar{y}$ : Media de los valores reales de la variable dependiente.
- $n$ : Número de observaciones.

#### - Interpretación

- $R^2 = 1$ : El modelo explica perfectamente toda la varianza de los datos.
- $R^2 = 0$ : El modelo no explica ninguna de la varianza; la predicción es equivalente a la media.
- $R^2 = <0$ : El modelo es peor que un modelo que simplemente predice la media de los valores observados.
- **Utilidad**

$R^2$  es útil para comprender qué tan bien el modelo se ajusta a los datos, aunque no indica necesariamente que el modelo sea bueno para hacer predicciones en nuevos datos.

#### 3.1.2. MAE (Error Absoluto Medio)

Mide el error promedio entre las predicciones del modelo y los valores reales. Se calcula como el promedio de las diferencias absolutas entre los valores reales y los valores predichos. Es una métrica simple e interpretativa, que refleja el tamaño promedio de los errores en las unidades de la variable objetivo (Mirajlili y Raschka, 2020). Donde:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- $y_i$ : Valor real.
- $\hat{y}_i$ : Valor predicho por el modelo.
- $n$ : Número de observaciones.

- **Interpretación**

- **MAE = 0**: No hay error en las predicciones (el modelo predice perfectamente).
- Cuanto mayor sea el **MAE**, peor es el rendimiento del modelo.

- **Utilidad**

**MAE** es útil para medir el error medio de predicción, es intuitivo y fácil de interpretar porque representa el error promedio en las mismas unidades que los datos.

### 3.1.3. RMSE (Raíz del Error Cuadrático Medio)

**RMSE** es similar al MAE, pero penaliza más los errores grandes. Es la raíz cuadrada del promedio de los cuadrados de las diferencias entre los valores reales y los valores predichos. Es una métrica más sensible a los outliers (valores atípicos) que el MAE (Mirajlili y Raschka, 2020). Donde:



$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

- $y_i$ : Valor real.
- $\hat{y}_i$ : Valor predicho por el modelo.
- $n$ : Número de observaciones.



#### - Interpretación

- **RMSE = 0**: No hay error en las predicciones (predicciones perfectas).
- Cuanto mayor sea el **RMSE**, peor es el rendimiento del modelo.
- Debido a que se eleva al cuadrado las diferencias antes de promediarlas, los errores más grandes tienen un impacto desproporcionado en el valor final de **RMSE**.

#### - Utilidad

**RMSE** es útil para medir la magnitud de los errores, penalizando más los errores grandes. Debe considerarse su uso preferentemente cuando los grandes errores son particularmente indeseables.

## CAPÍTULO IV

### RESULTADOS Y DISCUSIÓN

#### 4.1. DIFERENTES MODELOS DE APRENDIZAJE AUTOMÁTICO PARA LA ESTIMACIÓN DE LA MEDULACIÓN TOTAL EN FIBRA DE ALPACA HUACAYA BLANCO

Para el entrenamiento de los modelos de aprendizaje automático (AA), se realizó previamente un análisis exploratorio de datos (AED). En la tabla 1, se presenta las estadísticas descriptivas de las variables predictoras y la variable objetivo para 439 datos válidos.

**Tabla 1**

*Estadísticas Descriptivas para 04 características físicas de fibra*

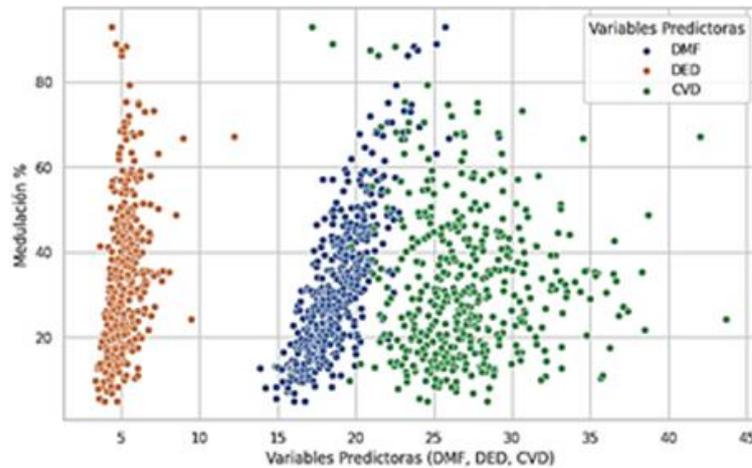
	N	Media	EE	DE	Mínimo	Máximo
DMF ( $\mu$ )	439	18.83	0.09	2.02	13.88	29.17
DED ( $\mu$ )	439	5.11	0.04	0.92	3.39	12.25
CVD (%)	439	27.16	0.17	3.70	17.20	43.65
MED (%)	439	32.16	0.80	16.89	4.91	92.87

DMF=Diámetro Medio de Fibra, DED= Desviación Estándar del Diámetro, CVD= Coeficiente de Variación del Diámetro  
MED= Medulación

El análisis de la fibra reveló un diámetro medio (DMF) de  $18.83 \pm 2.02$  micras, una desviación estándar del diámetro (DED) de  $5.11 \pm 0.92$  micras y un coeficiente de variación del diámetro (CVD) de  $27.16 \pm 3.70\%$ . La medulación total (MED) presentó una media de  $32.16\% \pm 16.89\%$ . Estos resultados difieren de los reportados por Porto (2017), quien registró un DMF de 20.90 micras, DED de 5.02 micras y CVD de 24.07%. Por su parte, Mamani P. (2024) documentó en Corani, tasas de medulación total diferenciadas por sexo, con 29.71% en hembras y 33.96% en machos.

## Figura 7

*Gráfico de dispersión de la relación entre variables predictoras y medulación.*



En la figura 07, se observa un gráfico de dispersión de los datos analizados, donde se visualiza la relación entre las variables predictoras (DMF, DED, y CVD) y la variable objetivo (MED), nótese que el DMF y la DED tiene una tendencia lineal, lo que no se observa con el CVD. La tabla 02, detalla las estadísticas descriptivas incluyendo el “Fenotipo” (Alta, Media y Baja). Se considera “Alta” cuando se observa en el animal la presencia de pelos en la nuca, pecho, espalda y vientre de la alpaca, “Media”, cuando la presencia de pelos solo es observable en vientre y pecho y “Baja”, cuando solo existe presencia de pelos en una de las 4 regiones.

**Tabla 2***Estadísticas Descriptivas entre Fenotipos y 04 características de fibra*

	<b>Fenotipo</b>	<b>N</b>	<b>Media</b>	<b>EE</b>	<b>DE</b>	<b>Mínimo</b>	<b>Máximo</b>
DMF	Alta	220	19.51	0.14	2.21	13.88	29.17
	Baja	68	17.75	0.14	1.20	14.89	22.06
	Media	151	18.34	0.13	1.63	14.22	22.72
DED	Alta	220	5.34	0.06	0.92	3.50	12.25
	Baja	68	4.63	0.07	0.64	3.39	6.31
	Media	151	5.00	0.07	0.93	3.54	9.49
CVD	Alta	220	27.43	0.24	3.68	17.20	41.99
	Baja	68	26.11	0.38	3.18	19.61	35.78
	Media	151	27.25	0.31	3.88	19.71	43.65
MED	Alta	220	42.81	1.11	16.49	11.11	92.87
	Baja	68	14.04	0.43	3.59	4.91	19.88
	Media	151	24.81	0.63	7.78	7.27	49.45

DMF=Diámetro Medio de Fibra, DED= Desviación Estándar del Diámetro, CVD= Coeficiente de Variación del Diámetro  
MED= Medulación

La tabla 2, hace notar que los fenotipos con Alta presencia de pelos, se relaciona con diámetros más altos, en tanto que la desviación estándar del diámetro tiene la misma tendencia, es decir a más alta presencia de pelos mayor será la desviación estándar del diámetro. Por otro lado, la relación de los promedios del coeficiente de variación del diámetro frente a los fenotipos, no guarda una relación semejante a la encontrada en el DMF y la DED. Pinares y Quispe (2024), estimaron una media de 32.78% de medulación en alpacas tui mayor procedentes del Cusco, en tanto que, Mamani R. (2023), determinó una medulación de 54.7 y 34.5 % en alpacas Huacaya y Suri respectivamente, procedentes de comunidades alpaqueras del distrito de San José – Azángaro – Puno, precisando que las alpacas son mayores a 5 años (Boca llena).

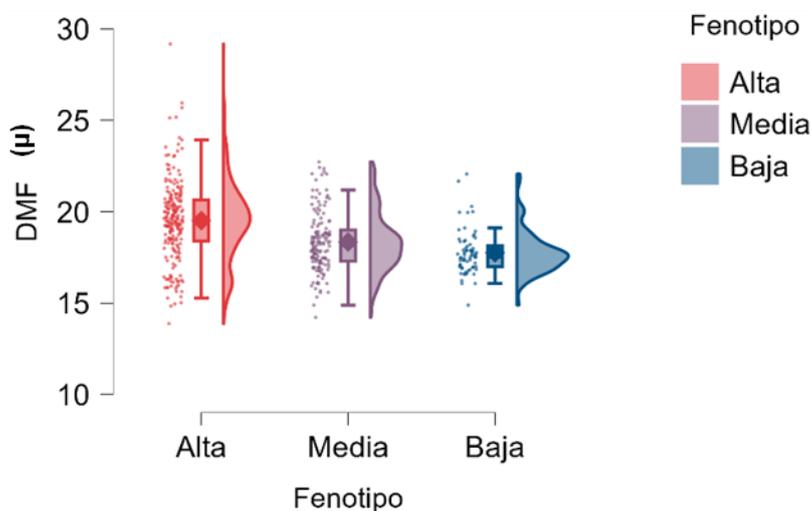
Por su parte, Mamani P. (2024), halló una medulación de 32.72, 34.98 y 33.95% en alpacas Huacaya, mayores a 4.5 años de edad (boca llena) en las comunidades de Chacaconiza, Quelcaya y Chimboya del distrito de Corani – Carabaya – Puno. Radzik-

Rant, y Wiercińska, (2021), determinaron un porcentaje de medulación en la lana de las alpacas polacas de 68.91%, siendo más alto en machos que en las hembras.

Por otro lado, la relación encontrada entre la medulación y los fenotipos hace notar que cuando es alta la presencia de pelos en el fenotipo, también será mayor la medulación tal como se observa en la figura 8.

### Figura 8

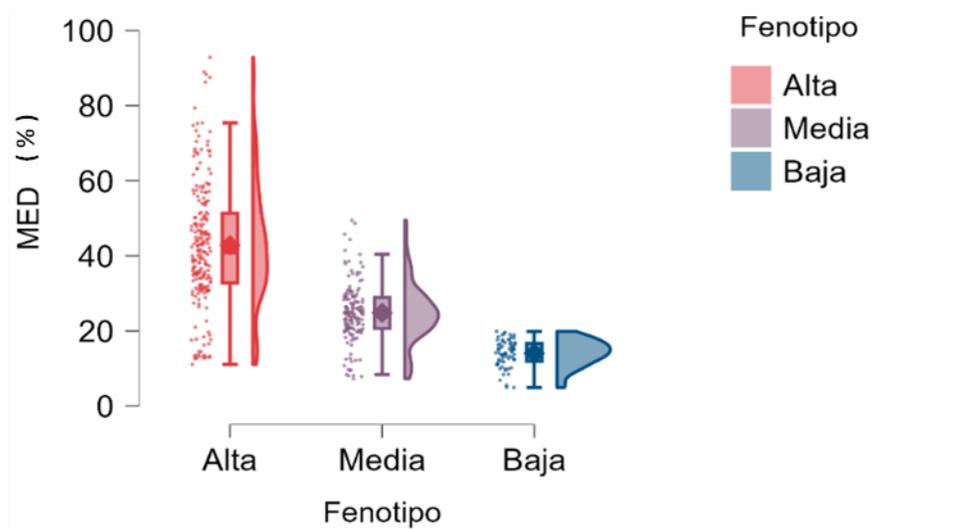
*Ploteo de la relación entre el DMF y los fenotipos.*



En tanto que, en la figura 9, se observa la misma tendencia entre los fenotipos y la Medulación, Es decir mientras sea alto el fenotipo también será alta la medulación.

### Figura 9

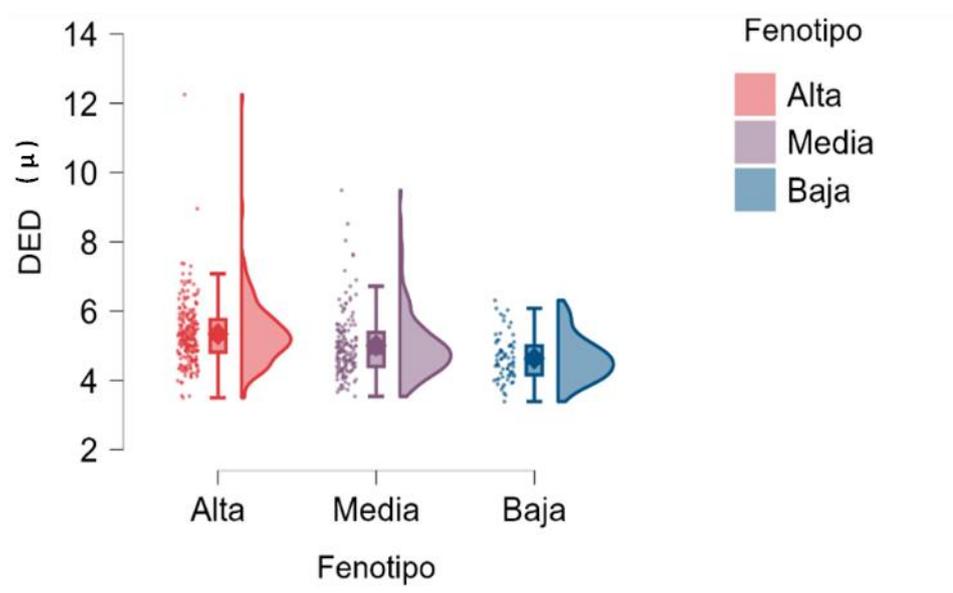
Ploteo de la relación entre la Medulación y los fenotipos.



La figura 10, muestra la relación entre la DED y los fenotipos Alto, Medio y Bajo, donde se observa que cuanto más presencia de pelos se observe en el fenotipo (Alto) también incrementará la DED.

### Figura 10

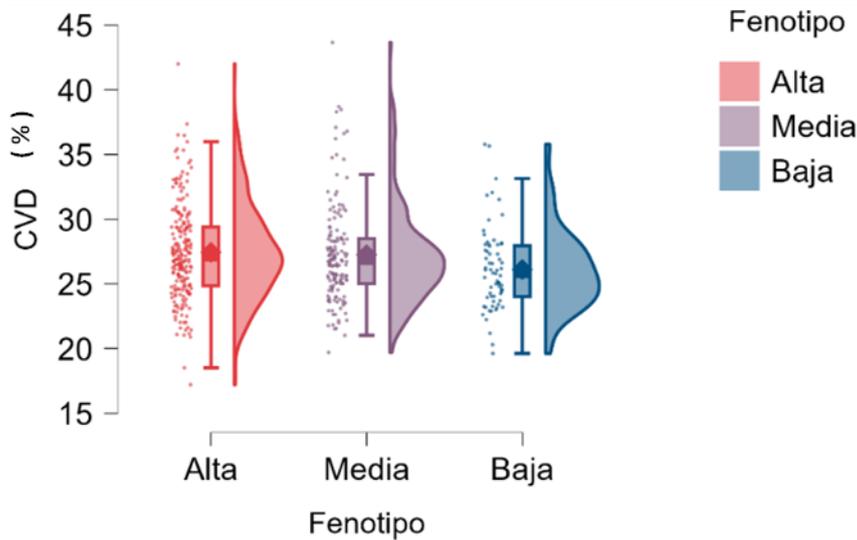
Ploteo de la relación entre la DED y los fenotipos.



Respecto a la relación entre los fenotipos y el CVD, no se observa una diferencia marcada entre los fenotipos alto, medio y bajo con el porcentaje del CVD.

**Figura 11**

*Ploteo de la relación entre el CVD y los fenotipos.*



La tabla 3, muestra una matriz de correlaciones entre cuatro variables: DMF (Diámetro Medio de Fibra), MED (Medulación), DED (Diámetro Estándar de la Fibra), y CVD (Coeficiente de Variación del Diámetro). Los coeficientes de correlación miden la fuerza y dirección de la relación lineal entre las variables.

- **DMF y MED** tienen una correlación fuerte positiva (0.76) y significativa ( $p < 0.001$ ).
- **DMF y DED** también están positivamente correlacionadas (0.59,  $p < 0.001$ ).
- **DED y CVD** muestran una correlación fuerte (0.74,  $p < 0.001$ ).
- Las demás correlaciones son bajas y no significativas. Esto sugiere que DMF, MED, y DED están fuertemente relacionadas, mientras que CVD no muestra una correlación fuerte con DMF o MED.

**Tabla 3***Matriz de Correlaciones entre variables predictoras y la variable objetivo*

	DMF ( $\mu$ )	MED (%)	DED ( $\mu$ )	CVD (%)
DMF ( $\mu$ )	—			
MED (%)	0.76 ***	—		
DED ( $\mu$ )	0.59 ***	0.53 ***	—	
CVD (%)	-0.05	0.04	0.74 ***	—

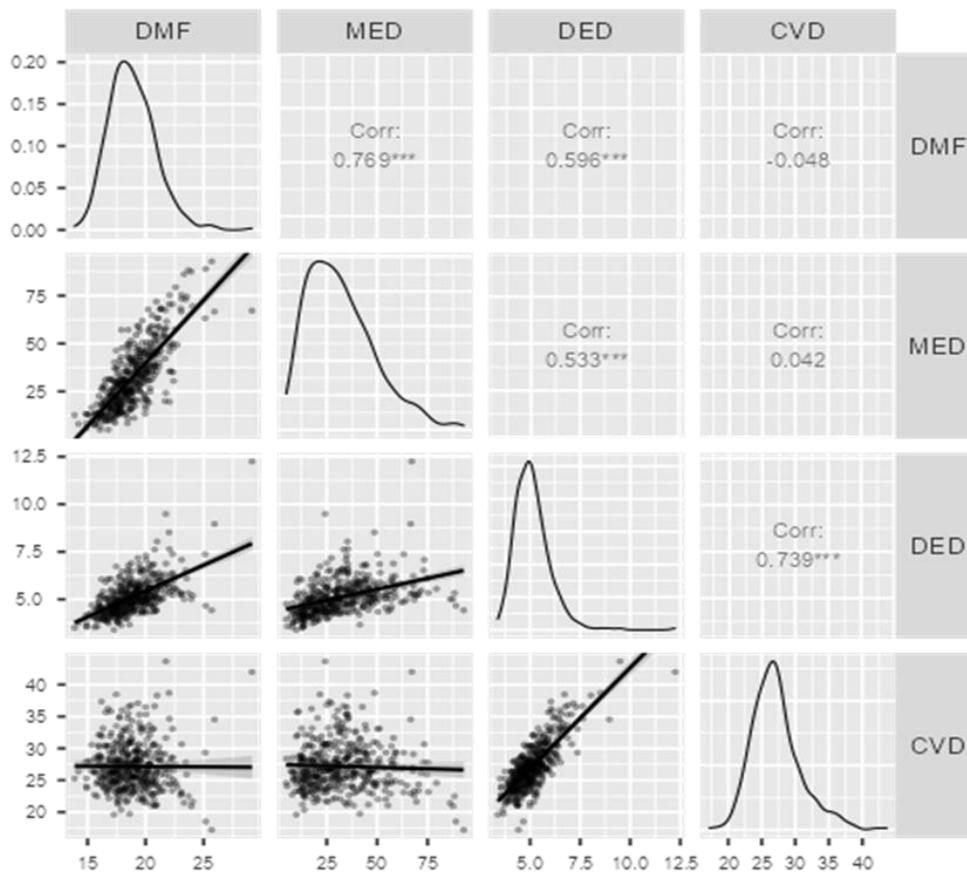
Nota. \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$

En contraste, Mamani P. (2024), estimó una correlación positiva moderada de 0.39 entre el DMF y MED en alpacas estudiadas en el distrito de Corani. Por su parte Porto (2017) estimó una correlación positiva moderada a alta de 0.64 entre el DMF y la DED y una correlación negativa baja de -0.15 entre el DMF y el CVD y una correlación positiva de moderada - alta de 0.61 entre la DED y el CVD ( $p < 0.01$ ), en su estudio de mapeo de la calidad de fibra llevado a cabo en la región de Puno. Por su parte Mamani R. (2024), estimó una correlación significativa de 0.36 entre la DMF y MED (en fibras fuertemente meduladas).

En la figura 12, se observa un Correlograma, que incluye tanto histogramas de las variables individuales (en la diagonal) como gráficos de dispersión y coeficientes de correlación para las combinaciones de las variables DMF, MED, DED y CVD. Los gráficos en la diagonal (histogramas), muestran la distribución de cada variable. Los gráficos de dispersión, revelan las relaciones entre pares de variables, como DMF y MED, que presentan una correlación positiva y fuerte (0.769\*\*\*), o entre DED y CVD, que muestran una correlación menor. Los asteriscos en los coeficientes indican la significancia estadística ( $p < 0.001$ ).

**Figura 12**

*Correlograma de Spearman para variables predictoras y variable objetivo.*



#### **4.2. RENDIMIENTO DE LOS ALGORITMOS DE APRENDIZAJE AUTOMÁTICO, EN LA ESTIMACIÓN DE LA TASA DE MEDULACIÓN TOTAL EN LA FIBRA DE ALPACA HUACAYA BLANCO.**

La tabla 4, presenta una comparación de 07 modelos, basado en el coeficiente de determinación ( $R^2$ ), el cual mide qué tan bien un modelo puede explicar la variabilidad de los datos.

**Tabla 4**

*Comparación de modelos usando el coeficiente de determinación*

	kNN	Gradient Boosting	Random Forest	Linear Regression	AdaBoost	Neural Network	Tree
kNN		0.04	0.02	0.03	0.05	0.54	0.03
Gradient Boosting	0.96		0.21	0.08	0.15	0.09	0.06
Random Forest	0.97	0.78		0.03	0.03	0.08	0.02
Linear Regression	0.97	0.91	0.96		0.33	0.11	0.03
AdaBoost	0.94	0.84	0.96	0.66		0.16	0.22
Neural Network	0.45	0.90	0.91	0.89	0.83		0.84
Tree	0.96	0.94	0.97	0.96	0.77	0.15	

Nota: La tabla muestra las probabilidades de que la puntuación para el modelo en la fila sea mayor que la del modelo en la columna

La tabla muestra las probabilidades de que el coeficiente de determinación de un modelo (fila) sea superior al de otro modelo (columna). Los valores más pequeños indican que el modelo de la fila tiene una mayor probabilidad de superar al modelo de la columna. Un valor cercano a 1 indica que no hay diferencias significativas en el desempeño entre ambos modelos. A continuación, se presenta las principales observaciones por modelo:

**a. kNN**

- Se desempeña peor que Gradient Boosting y Random Forest, como lo indican los valores bajos (0.04 y 0.02 respectivamente).
- Supera ligeramente a AdaBoost (0.05) y se compara favorablemente frente a regresión lineal (0.03), aunque estas diferencias no son muy grandes.

**b. Gradient Boosting**

- Funciona extremadamente bien, superando a la mayoría de los modelos, como se observa en las comparaciones frente a regresión lineal (0.08) y Neural Network (0.09).



- Es comparable al modelo Random Forest (0.78), lo que sugiere que ambos modelos tienen un rendimiento similar.

#### **c. Random Forest**

- Presenta un rendimiento sólido, superior al de muchos modelos (como regresión lineal y Neural Network).
- Tiene probabilidades cercanas a 1 en varias comparaciones, lo que sugiere que sus diferencias con otros modelos son insignificantes (por ejemplo, frente a Gradient Boosting y AdaBoost).

#### **d. Linear Regression**

- Aunque su desempeño es bastante bajo en general, parece tener mejores resultados frente a modelos como kNN y AdaBoost.
- Los valores de probabilidad más bajos se encuentran en las comparaciones con modelos basados en árboles como Gradient Boosting y Random Forest.

#### **e. AdaBoost**

- No se desempeña tan bien como Gradient Boosting y Random Forest, pero supera ligeramente a kNN y es comparable a regresión lineal en cuanto a su ajuste.

#### **f. Neural Network**

- Se desempeña peor en comparación con modelos como Gradient Boosting, Random Forest y regresión lineal.
- Su probabilidad de superarlos es baja (0.09, 0.08), lo que indica un rendimiento inferior.



**g. Tree**

- Funciona de manera similar a otros modelos basados en árboles, pero tiene una probabilidad relativamente baja frente a Gradient Boosting (0.06).
- Su desempeño frente a Neural Network es mejor (0.84), lo que sugiere que tiene un rendimiento superior.

Los resultados evidencian que los modelos basados en ensambles de árboles, particularmente Random Forest y Gradient Boosting, mostraron un desempeño superior en la predicción de la medulación de fibra de alpaca, superando significativamente a modelos más simples como kNN, regresión lineal y redes neuronales. Random Forest destaca como el modelo más eficiente y preciso.

**4.3. EFICACIA DEL MODELO UTILIZANDO UN CONJUNTO DE DATOS INDEPENDIENTES MEDIANTE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA LA ESTIMACIÓN DE LA MEDULACIÓN EN FIBRA DE ALPACA HUACAYA BLANCO.**

En la tabla 5, se observa los resultados de las métricas de desempeño en la predicción de la medulación, donde el mejor valor de  $R^2$  lo tiene el modelo Random Forest con un valor de 0.85. Esto indica que el modelo explica el 85% de la variabilidad de los datos. Le siguen Gradient Boosting (0.83) y Linear Regression (0.81).

Respecto al valor más bajo de RMSE, que indica el error promedio más pequeño, también corresponde a Random Forest con 6.53; el segundo mejor valor lo tiene Gradient Boosting con 6.82. Por otro lado, el menor error absoluto medio (MAE) lo tiene nuevamente Random Forest con 4.77, en tanto que, el segundo lugar lo obtiene Gradient Boosting con 4.96.

**Tabla 5***Métricas de Desempeño para 7 modelos predictores de Medulación*

N°	Modelo	Train	Test	R <sup>2</sup>	RMSE	MAE
1	kNN	0.02	0.02	0.71	8.98	6.70
2	Gradient Boosting	0.33	0.00	0.83	6.82	4.96
3	Random Forest	0.29	0.03	0.85	6.53	4.77
4	Linear Regression	0.03	0.01	0.81	7.34	5.66
5	AdaBoost	0.37	0.01	0.80	7.50	5.25
6	Neural Network	1.43	0.01	0.71	9.08	6.86
7	Tree	0.38	0.00	0.78	7.80	5.71

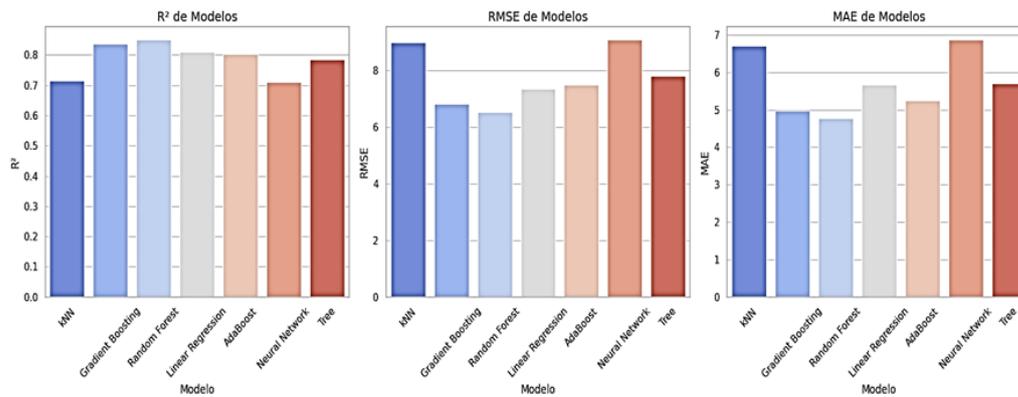
Donde: R<sup>2</sup> = Coeficiente de Determinación, RMSE = Error Cuadrático Medio y MAE = Error Medio Absoluto

De acuerdo a las métricas expuestas: El modelo Random Forest es el más eficiente, ya que tiene el mayor valor de R<sup>2</sup> (0.85), el menor RMSE (6.53) y el menor MAE (4.77). Esto sugiere que Random Forest tiene un mejor ajuste y produce errores menores en comparación con los otros modelos. Gradient Boosting es el segundo mejor, con valores de R<sup>2</sup> y errores cercanos a los de Random Forest, pero ligeramente con mayores tasas de error en RMSE y MAE.

En resumen, observando la figura 13, el modelo Random Forest tiene un balance favorable entre las tres métricas clave, lo que lo convierte en el modelo más eficiente y robusto para la predicción de la medulación en el presente estudio.

**Figura 13**

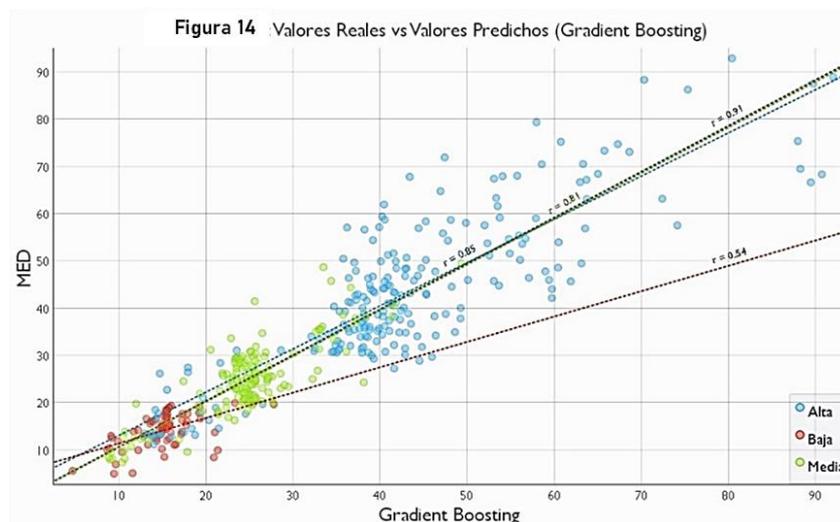
*Gráfico de barras comparativo para 3 métricas de desempeño de AA.*



El modelo Gradient Boosting (Figura 14), indica un rendimiento bastante bueno para los fenotipos con valores más altos en contraste con la medulación, como "Alta" y "Media", pero tiene un desempeño moderado con el fenotipo "Baja". Si se busca mejorar el rendimiento para esta última categoría, se podrían considerar ajustes en el modelo o un análisis más detallado de las características del fenotipo "Baja". El alto valor de "r" para los otros dos grupos muestra que el modelo captura bastante bien las tendencias generales de los datos.

**Figura 14**

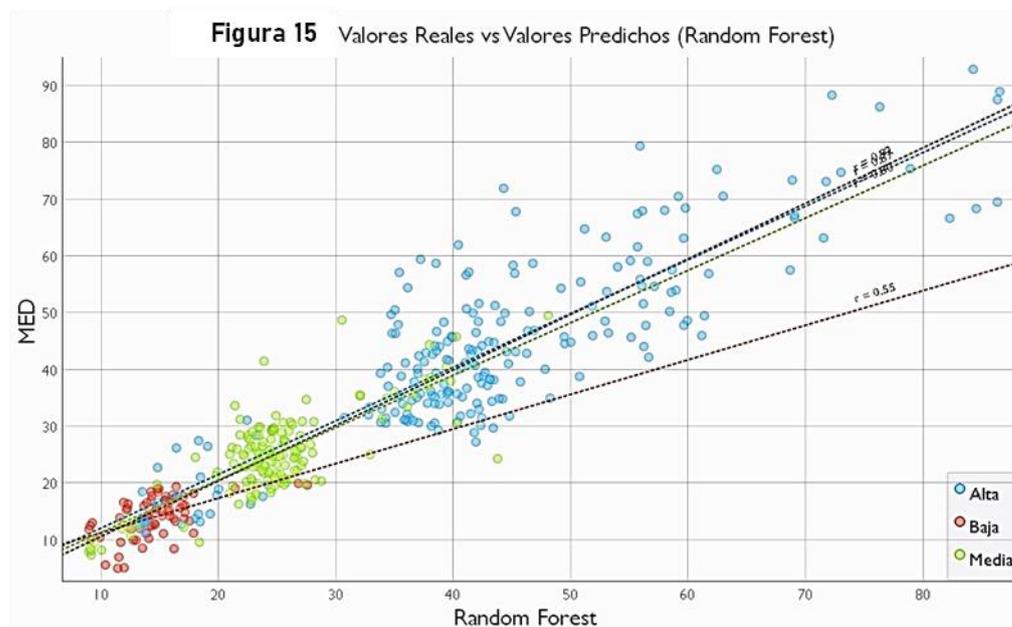
*Valores reales vs valores predichos*



El modelo Random Forest (figura 15), es altamente preciso para los fenotipos con valores más altos de Medulación, especialmente el fenotipo "Alta" y, en menor medida, el fenotipo "Media". Sin embargo, tiene problemas para ajustar adecuadamente el fenotipo "Baja", lo que sugiere que se podría investigar la posibilidad de ajustar el modelo o utilizar otras técnicas para mejorar el rendimiento en ese grupo específico.

### Figura 15

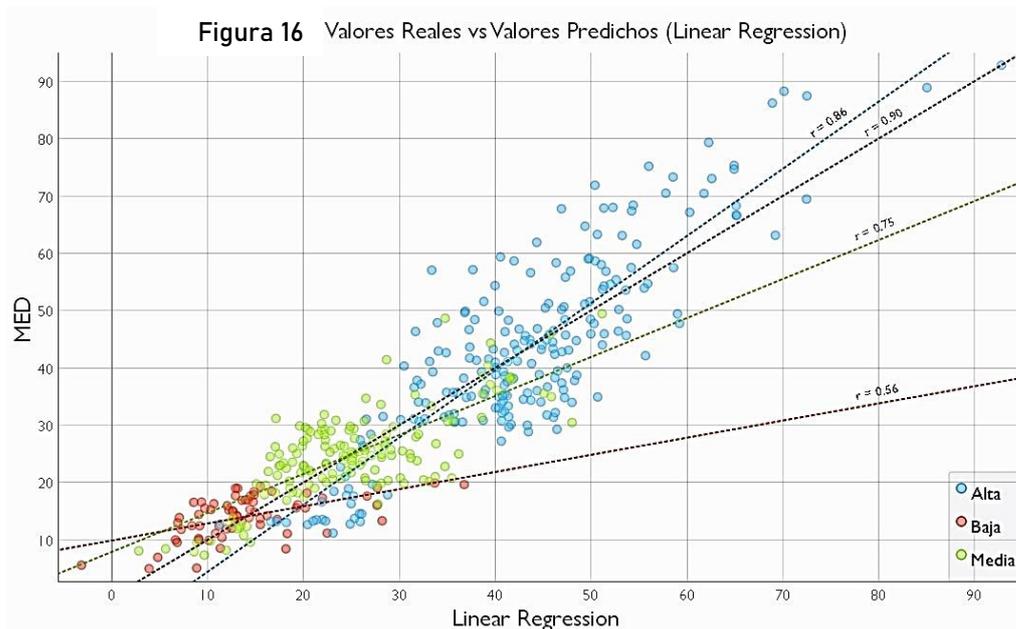
*Valores reales vs valores predichos (Random Forest)*



Por otra parte, la figura 16 sugiere que hay una relación lineal entre los valores reales y los predichos, con diferentes grados de correlación según la categoría. Esto indica que la precisión de la predicción varía dependiendo de si los valores son bajos, medios o altos.

**Figura 16**

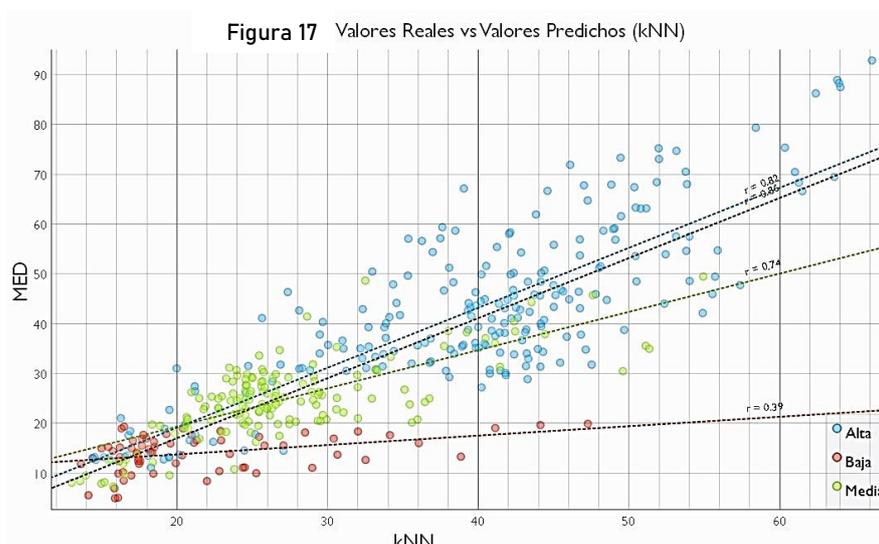
*Valores reales vs valores predichos linear regresión*



Respecto a la figura 17, se visualiza cómo se comparan los valores predichos por el modelo kNN con los valores reales, lo cual es útil para evaluar el rendimiento del modelo predictivo. La correlación moderada indica que el modelo tiene una precisión aceptable, pero hay margen para mejorar.

**Figura 17**

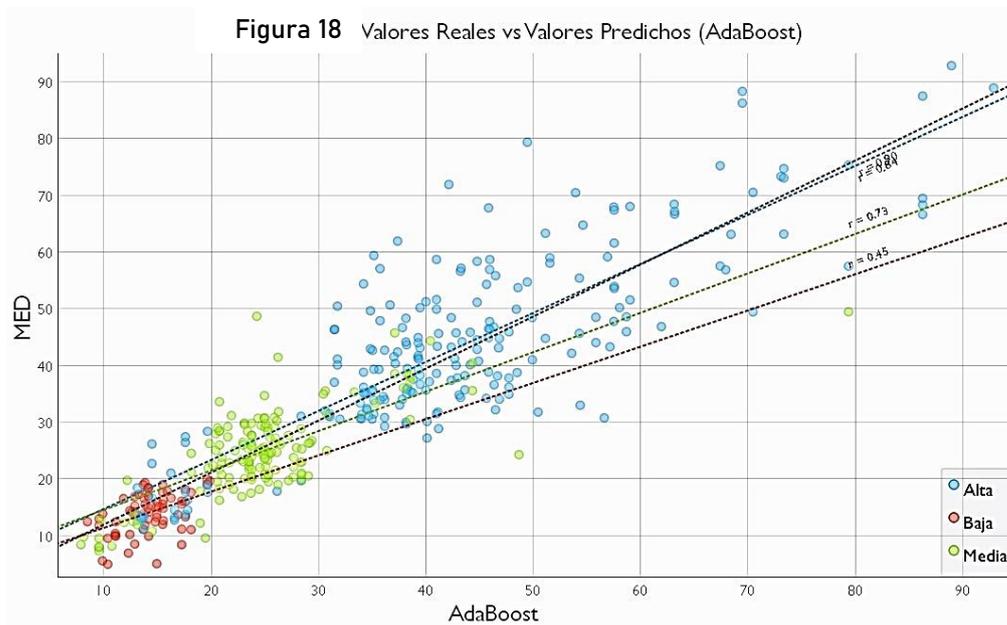
*Valores reales vs valores predichos kNN*



La figura 18 sugiere que el modelo AdaBoost tiene un rendimiento moderado en la predicción de valores reales (MED). La precisión varía según la categoría de los datos, siendo más precisa en los valores medios y menos en los valores altos y bajos. Las líneas de error indican que hay una variabilidad en las predicciones, y el coeficiente de correlación moderado sugiere que hay espacio para mejorar el modelo

### Figura 18

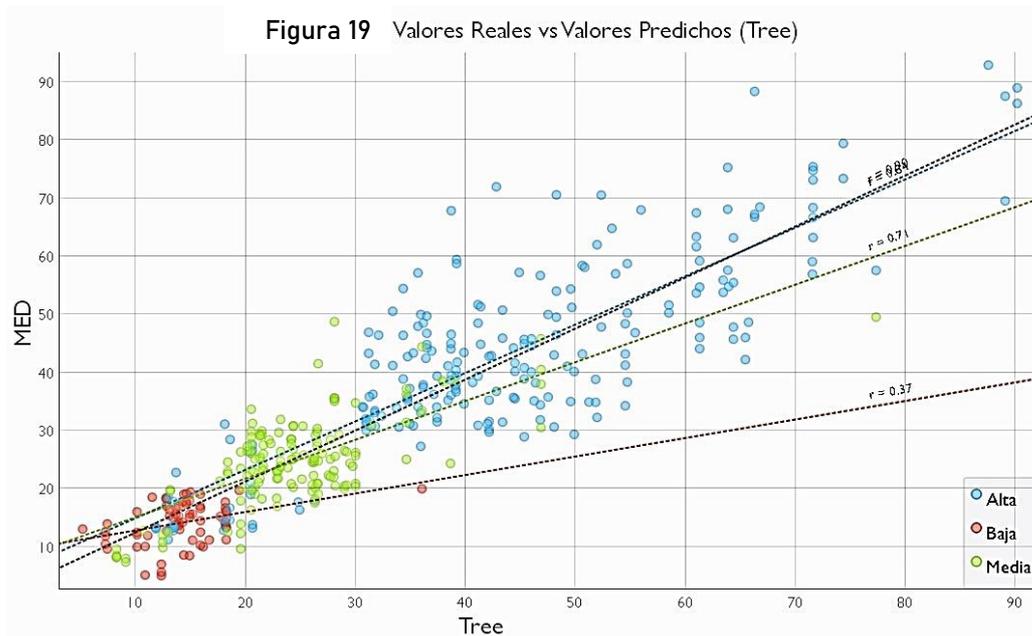
*Valores reales vs valores predichos AdaBoost*



La figura 19 sugiere que el modelo de árbol de decisión tiene un rendimiento razonable, especialmente para valores altos, pero hay margen de mejora, particularmente en las predicciones de valores medios y bajos.

## Figura 19

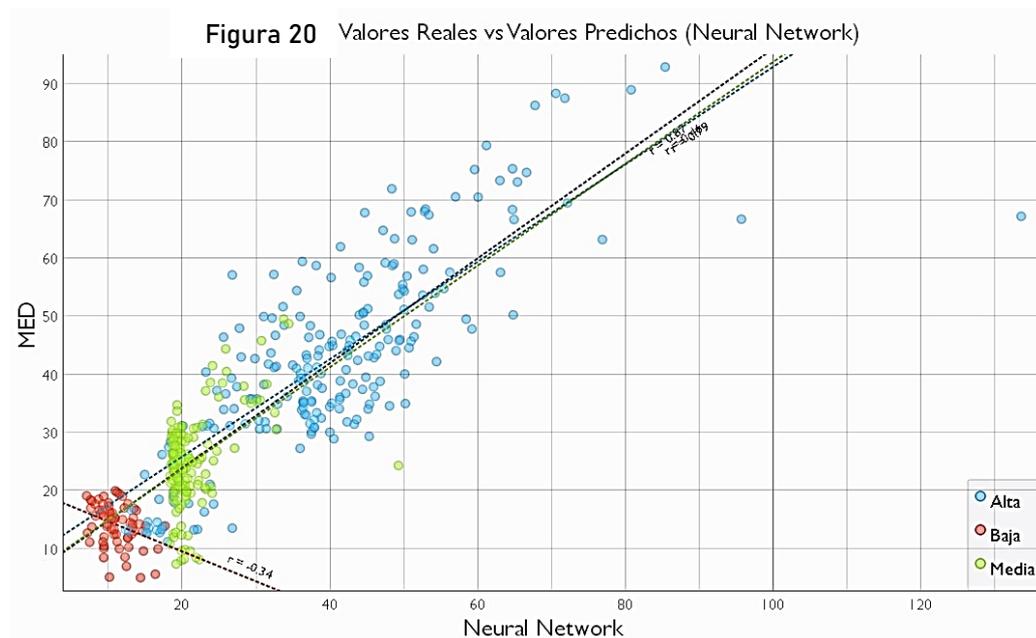
Valores reales vs valores predichos Tree



El modelo de red neuronal (Figura 20), parece funcionar bien para los fenotipos con valores más altos de Medulación (fenotipo "Alta" y en parte "Media"), pero tiene problemas significativos para los valores bajos, como se observa en el fenotipo "Baja". En líneas generales el modelo indica un buen ajuste global, pero hay áreas específicas donde el modelo no capta la variabilidad de los datos correctamente.

## Figura 20

Valores reales vs valores predichos Neural Network



### 4.4. DISCUSIÓN

Estudios similares como el de Quispe et al. (2022), emplearon modelos de regresión lineal y regresión múltiple en la predicción de la medulación, encontrando que el DMF logra explicar el 70% de la variabilidad en la medulación a través del uso del dispositivo denominado FIBERMED. Bonilla et al., (2023) realizó análisis de los coeficientes de correlación entre los métodos de evaluación de medulación y demostró una alta robustez estadística, con valores de  $r = 0.99$  y  $r = 0.97$  para las comparaciones entre el método basado en Inteligencia Artificial y Análisis de imágenes digitales respectivamente, frente a el método de microscopio de proyección tradicional.

Quispe et al. (2023), realizó el análisis comparativo de los métodos de medición de medulación en fibras de alpaca y reveló diferencias significativas entre los sistemas evaluados. El OFDA100 registró consistentemente valores inferiores de medulación ( $9.81 \pm 0.61\%$  a  $18.30 \pm 1.68\%$ ) en comparación con el microscopio de proyección tradicional



( $67.43 \pm 1.74\%$ ) y el microscopio de proyección computarizado ( $28.71 \pm 1.37\%$  a  $41.14 \pm 2.00\%$ ). Estos resultados sugieren que el OFDA100, a pesar de su eficiencia en la evaluación de medulación en lanas, presenta limitaciones significativas para la caracterización precisa de la medulación en fibras de alpaca, evidenciando una subestimación sistemática de los valores de medulación total.

Pinares et al. (2018), en su estudio reveló hallazgos significativos respecto a la medulación en fibras de alpaca, evidenciando un rango de variación del porcentaje de fibras meduladas entre  $12.33\%$  y  $91.67\%$  por muestra. Los análisis genéticos, basados en modelos univariados y bivariados, estimaron una heredabilidad máxima para la categoría de medulación (CM) de  $0.36 \pm 0.13$ , mientras que para el diámetro de fibra (FD) fue de  $0.35 \pm 0.15$ , con una correlación genética significativa entre ambos caracteres de  $0.93 \pm 0.12$ .

Shahinfar y Kahn (2018), evaluaron el rendimiento de modelos de aprendizaje automático, utilizaron métricas como el error absoluto medio (MAE) y el coeficiente de correlación y los resultados mostraron que el método MT (Model Tree) obtuvo los mejores valores, con coeficientes de correlación de 0.93, 0.90, 0.94, 0.81 y 0.59 para predecir aGFW (peso de lana sucia), aCFW (peso de lana limpia), aFD (diámetro de fibra), aSL (Longitud de fibra) y aSS (Resistencia de la fibra), respectivamente, y un MAE de 0.48 kg, 0.41 kg, 0.92  $\mu\text{m}$ , 6.91 mm y 6.82 N/ktex. Estos resultados sugieren que el método MT es una herramienta prometedora para la predicción temprana de características de la lana en ovejas adultas.

Sabuncu y Özdemir (2023), demostraron que la técnica de tomografía de coherencia óptica espectral (OCT) combinada con algoritmos de aprendizaje automático



puede ser aplicada de manera efectiva para distinguir la lana Merino de la lana gruesa. Este método no destructivo y sin contacto obtuvo una precisión de 0.946.

Yildiz, K. (2020), en su investigación presenta un nuevo método para distinguir las fibras de lana y mohair utilizando el análisis de imágenes, empleando patrones binarios locales y técnicas de aprendizaje profundo para analizar imágenes microscópicas de estas fibras. Los resultados demostraron altas tasas de precisión, con el aprendizaje profundo alcanzando el 99.8% y el aprendizaje automático el 90.25% en la identificación correcta de lana y mohair. Este enfoque ofrece una alternativa más objetiva, eficiente y rentable a los métodos manuales tradicionales utilizados en la industria textil.

Xing et al. (2022), en su investigación presenta un nuevo método altamente preciso para distinguir entre fibras de lana y cachemira. Utilizando técnicas de aprendizaje profundo, específicamente redes neuronales convolucionales pre-entrenadas (como VGG-16), los investigadores lograron identificar con una exactitud del 99.15% diferentes tipos de fibras a partir de imágenes microscópicas. Esta metodología representa un avance significativo en comparación con los métodos tradicionales de aprendizaje automático, ofreciendo una solución más eficiente y precisa para la identificación de fibras en la industria textil.

Manivannan et al. (2022), en su trabajo de investigación evalúa un algoritmo híbrido para optimizar y predecir parámetros del mecanizado por chorro de agua abrasivo (AWJM) en composites de poliéster reforzado con lana de oveja. El método MOORA identificó las variables más influyentes, y el modelo híbrido MOORA—SVM—Entropy mejoró el rendimiento de predicción en un 38.7% en comparación con el modelo SVM estándar.



Mancisidor et al. (2021), en sus resultados para el porcentaje de medulación (PM) muestran que la precisión de las predicciones genómicas mejoró en un 1.471% en promedio en comparación con el método BLUP clásico. Por tanto, sugieren que la incorporación de datos genómicos puede contribuir a reducir la medulación en alpacas Huacaya, un rasgo importante para la calidad de la fibra.



## V. CONCLUSIONES

- El entrenamiento de los modelos de aprendizaje automático fue satisfactorio, teniendo a las variables predictoras; Diámetro Medio de fibra (DMF), Desviación Estándar del Diámetro (DED) y Coeficiente de Variación del Diámetro (CVD) con los siguientes promedios;  $18.83 \pm 2.02$  micras,  $5.11 \pm 0.92$  micras y  $27.16 \pm 3.70$  % respectivamente, en tanto que, la variable objetivo, Medulación (MED) tuvo una media de  $32.16 \pm 16.89$  %.
- La comparación de modelos imputó a Random Forest como el modelo más robusto y consistente, demostrando una probabilidad superior al 96% de superar a la mayoría de los modelos evaluados.
- El modelo que se validó como mejor estimador de la medulación fue “**Random Forest**”, logrando explicar el 85% de la variabilidad de la Medulación en la fibra de alpaca.



## VI. RECOMENDACIONES

- Se recomienda ampliar el número de muestras en futuros estudios para contar con un dataset más robusto, para mejorar el rendimiento de los modelos entrenados.
- Se recomienda también a partir de la presente investigación, incorporar muestras de alpacas Huacaya de diferentes edades a fin de contar con un modelo predictivo integral para alpacas Huacaya de color de vellón blanco.
- Para mejorar el rendimiento del modelo validado en el presente estudio, se sugiere hacer ajustes de los parámetros del modelo Random Forest, ampliando los datos de entrenamiento a 80% (datos de entrenamiento) y 20% (prueba) para las pruebas de validación del modelo.



## VII. REFERENCIAS BIBLIOGRÁFICAS

- Aylan-Parker, J., & McGregor, B. A. (2002). Optimising sampling techniques and estimating sampling variance of fleece quality attributes in alpacas. *Small Ruminant Research*, 44(1), 53-64.
- Balasingam, A. (2005). The definitions of medullation threshold values used by different testing methods to define an objectionable medullated fibre in Merino Wool. *Aust. Wool Innovation Proj.*
- Balasingam, A., & Mahar, T. J. (2005.). Status report on dark and medullated fibre testing of presale core samples and review of the detection threshold for contaminant medullation. In *International Wool Textile Organisation, Technology and Standards Committee, Hobart Meeting Report No. RWG (Vol.4)*.
- Bonilla, M. Q., Serrano-Arriezu, L., Trigo, J. D., Bonilla, C. Q., Gutiérrez, A. P., & Pena, E. Q. (2022). Application of artificial intelligence and digital images analysis to automatically determine the percentage of fiber medullation in alpaca fleece samples. *Small Ruminant Research*, 213, 106724.
- Carpio, M. (1991). La fibra de camélidos. *Producción de rumiantes menores: Alpacas. Lima: RERUMEN. p, 297-359.*
- Chattopadhyay, S. K., Bhaskar, P., Ahmed, M., Gupta, N. P., & Pokharna, A. K. (2005). Properties of indigenous angora rabbit hair and cotton blended yarns spun using shortstaple cotton spinning system.
- Contreras, A. (2010). *Estructura y características físicas de la fibra de alpaca Huacaya (Vicugna pacos) de color blanco* (Doctoral dissertation, Tesis de Ingeniero Zootecnista. Huancavelica, Perú: Univ. Nacional de Huancavelica).
- Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2009). Introduction to algorithms (3rd ed.). The MIT Press.
- Cruz, A., Morante, R., Gutiérrez, J., Torres, R., Burgos, A., & Cervantes, I. (2019). Genetic parameters for medullated fiber and its relationship with other productive traits in alpacas.. *Animal : an international journal of animal bioscience*, 1-7 . <https://doi.org/10.1017/S1751731118003282>.



- Cruz, A., Morante, R., Gutiérrez, J., Torres, R., Burgos, A., & Cervantes, I. (2018). Genetic parameters for medullated fiber and its relationship with other productive traits in alpacas.. *Animal : an international journal of animal bioscience*, 1-7 .  
<https://doi.org/10.1017/S1751731118003282>.
- Czaplicki, Z. (2012). Properties and structure of polish alpaca wool. *Fibres Text. East. Eur*, 20, 8-12.  
*EC651, 32pp.*
- Frank, E. N., Hick, M. V. H., Molina, M. G., & Caruso, L. M. (2009). Heredabilidades, correlaciones fenotípicas, correlaciones genéticas y repetibilidades de variables productivas en llamas argentinas. *Archivos Latinoamericanos de Producción Animal*, 17(3-4), 103-107.
- Gupta, N. P., Arora, R. K., & Verma, G. K. (1981). An Assessment of the characteristics of Medullated and Non-medullated Wool Fibres.
- Hack, W., McGregor, B., Ponzoni, R., Judson, G., Carmichael, I., & Hubbard, D. (1999). *Australian alpaca fibre: improving productivity and marketing-a report for the Rural Industries Research and Development Corporation*. Deakin University.
- Huanca Mamani, T. (1996). Manual del Alpaquero. In *Manual*; Instituto Nacional de Investigación Agraria.
- Hunter, L. (1993). *Mohair: a review of its properties, processing and applications*.
- Hunter, L., Smuts, S., & Botha, A. F. (2013). Characterizing visually objectionable and nonobjectionable medullated fibers in mohair. *Journal of natural fibers*, 10(2), 112-135.
- Laime Huarcaya, F. D. M., Pinares Huamaní, R., Paucara Ocsa, V., Machaca Machaca, V., & Quispe Peña, E. C. (2016). Características tecnológicas de la fibra de llama (Lama glama) Chaku antes y después de descender. *Revista de Investigaciones Veterinarias del Perú*, 27(2), 209-217.
- Lupton, C. J., & Pfeiffer, F. A. (1998). Measurement of medullation in wool and mohair using an optical fibre diameter analyser. *Journal of Animal Science*, 76(5), 1261-1266.



- Mamani Mamani, R. (2023). *Tasa de medulación y su correlación con el diámetro de fibra en alpacas Huacaya y Suri en las comunidades del distrito de San José, Azángaro.*
- Mamani, P. (2024). Caracterización fenotípica de fibras meduladas y su correlación con el diámetro de fibra y el índice de curvatura en alpacas Huacaya del distrito de Corani. *Universidad Nacional del Altiplano - Puno.*
- Mancisidor, B., Cruz, A., Gutiérrez, G., Burgos, A., Morón, J. A., Wurzinger, M., & Gutiérrez, J. P. (2021). ssGBLUP method improves the accuracy of breeding value prediction in Huacaya Alpaca. *Animals, 11*(11), 3052.
- Manivannan, J., Rajesh, S., Mayandi, K., Abuthakeer, S. S., Ravichandran, M., Kumar, T. S. M., ... & Siengchin, S. (2022). A novel and prediction approach of sheep wool reinforced polyester composites: Surface qualities and hybrid modeling. *Polymer Composites, 43*(8), 5274-5290.
- Martinez, Z., Iñiguez, L. C., & Rodríguez, T. (1997). Influence of effects on quality traits and relationships between traits of the llama fleece. *Small Ruminant Research, 24*(3), 203-212.
- McGregor, B. A. (2006). Production, attributes and relative value of alpaca fleeces in southern Australia and implications for industry development. *Small Ruminant Research, 61*(2-3), 93-111.
- McGregor, B. A., Butler, K. L., & Ferguson, M. B. (2013). The relationship of the incidence of medullated fibres to the dimensional properties of Mohair over the lifetime of Angora goats. *Small ruminant research, 115*(1-3), 40-50.
- McGregor, B. A., Ramos, H. E., & Peña, E. Q. (2012). Variation of fibre characteristics among sampling sites for Huacaya alpaca fleeces from the High Andes. *Small Ruminant Research, 102*(2-3), 191-196.
- Montes, M., Quicaño, I., Quispe, R., Quispe, E., & Ruiz, L. A. (2008). Quality characteristics of Huacaya alpaca fibre produced in the Peruvian Andean Plateau region of 611 Huancavelica. *Spanish Journal of Agricultural Research, (1)*, 33-38.



- Moore, K. E. (2015). The impact of fleece characteristics on insulation and heat exchange, and the consequential effect on vitamin D of alpacas in southern Australia.
- Moore, K. E., Blache, D., & Maloney, S. K. (2011). Fibre diameter and insulation in alpacas: The biophysical implications. *Small ruminant research*, 96(2-3), 165-172.
- Núñez Reiz, A., Armengol de la Hoz, M. A., & Sánchez García, M. (2019). Big data analysis and machine learning in intensive care units. *Big Data Analysis and Machine Learning in Intensive Care Units*, 43(7), 416-426.
- Onal, L., Korkmaz, M., & Tutak, M. (2007). Relations between the characteristics of Angora rabbit fibre. *Fibers and Polymers*, 8, 198-204.
- Orange: data mining toolbox in Python. *the Journal of machine Learning research*, 14(1), 2349- 2353.
- Ferguson, M. B., McGregor, B. A., & Behrendt, R. (2012). Relationships between skin follicle characteristics and fibre properties of Suri and Huacaya alpacas and Peppin Merino sheep. *Animal production science*, 52(7), 442-447.
- Pertuz, C. M. P. (2022). *Aprendizaje automático y profundo en python*. Ra-Ma Editorial.
- Pinares, R., & Quispe, E. (2024). Características de las fibras meduladas en alpaca huacaya y suri de Perú. *Chilean journal of agricultural & animal sciences*, 40(2), 332-340.
- Pinares, R., Gutiérrez, G. A., Cruz, A., Burgos, A., & Gutiérrez, J. P. (2019). Variabilidad fenotípica del porcentaje de fibras meduladas en el vellón de alpaca Huacaya. *Revista de Investigaciones Veterinarias del Perú*, 30(2), 699-708.
- Pinares, R., Gutiérrez, G. A., Cruz, A., Morante, R., Cervantes, I., Burgos, A., & Gutiérrez, J. P. (2018). Heritability of individual fiber medullation in Peruvian alpacas. *Small Ruminant Research*, 165, 93-100.
- Pinares, R., Gutiérrez, G. A., Cruz, A., Morante, R., Cervantes, I., Burgos, A., & Gutiérrez, J. P. (2018). Heritability of individual fiber medullation in Peruvian alpacas. *Small Ruminant Research*, 165, 93-100.



- Poma Copa, G. F. (2018). Evaluación de las características físicas de la fibra de llama (Lama glama) a la primera esquila en la mancomunidad de municipios Aymaras sin Fronteras para su procesamiento en la industria textil. *Apthapi*, 4, 1275.
- Porto, H. (2017). *Mapeo de la calidad de fibra en la región Puno*. Proyecto Especial Camélidos Sudamericanos – Puno - Perú.
- Quispe Huanca, J. L. (2014). Caracterización fenotípica de llamas del tipo t'amphulli conservadas en condición in situ en las regiones de Quetena Grande-Potosí y Calientes-Cochabamba. *Revista Científica de Investigación INFO-INIAF*, 1, 54.
- Quispe, E. C., Alfonso, L., Flores, A., Guillén, H., & Ramos, Y. (2009). Bases para un programa de mejora de alpacas en la región altoandina de Huancavelica-Perú. *Archivos de zootecnia*, 58(224), 705-716.
- Quispe, E. C., Rodríguez, T. C., Iñiguez, L. R., & Mueller, J. P. (2009). Producción de fibra de alpaca, llama, vicuña y guanaco en Sudamérica. *Animal Genetic Resources/Recursos genéticos animales/Recursos genéticos animales*, 45, 1-14.
- Quispe, E. C., Torres, R., Pinares, R., Cruz, A., & Bonilla, M. Q. (2023). Comparison of projection microscope with OFDA100 in alpaca fibers medullation. *Archivos de zootecnia*, 72(279), 226-234.
- Quispe, E., Quispe, M., Quispe, C., Poma, A., Paucar-Chanca, R., Cruz, A., & McGregor, B. (2022). Relationships between the incidence and degree of medullation with the diameter of alpaca fibers evaluated using a novel device based on artificial intelligence. *The Journal of The Textile Institute*, 114, 1016 - 1031. <https://doi.org/10.1080/00405000.2022.2105110>.
- Quispe, M. D., Quispe, C. C., Serrano-Arriezu, L., Trigo, J. D., Bengoechea, J. J., & Quispe, E. C. (2023). Development and validation of a smart system for medullation and diameter assessment of alpaca, llama and mohair fibres. *animal*, 17(5), 100800.
- Radzik-Rant, A., & Wiercińska, K. (2021). Analysis of the wool thickness and medullation characteristics based on sex and color in a herd of alpacas in Poland. *Archives Animal Breeding*, 64(1), 157-165.



- Radzik-Rant, A., Pofelska, O., & Rant, W. (2018). Characteristics of alpaca wool from farmed animals located on different continents. *Annals of Warsaw University of Life Sciences-SGGW. Animal Science*, 57.
- Radzik-Rant, A., Wielechowska, M., & Rant, W. (2021). Variation in Wool Characteristics across the Body in a Herd of Alpacas Kept in Poland. *Animals : an Open Access Journal from MDPI*, 11. <https://doi.org/10.3390/ani11102939>.
- Rafat, S. A., De Rochambeau, H., Brims, M., Thébault, R. G., Deretz, S., Bonnet, M., & Allain, D. (2007). Characteristics of Angora rabbit fiber using optical fiber diameter analyzer. *Journal of animal science*, 85(11), 3116-3122.
- Raschka, Sebastian., & Mirjalili, Vahid. (2020). Python machine learning : machine learning and deep learning with Python, scikit-learn, and TensorFlow. Packt Publishing.
- Raschka, S., Liu, Y. H., & Mirjalili, V. (2022). *Machine Learning with PyTorch and Scikit-Learn: Develop machine learning and deep learning models with Python*. Packt Publishing Ltd.
- Rodríguez, T. (2003). Calidad de fibra de llama descordada y clasificada. *Uso pastoril en humedales altoandinos*, 93.
- Sabuncu, M., & Özdemir, H. (2023). Optical coherence tomography imaging can identify Merino lambs' wool using automatic machine learning vision. *Textile Research Journal*, 93(19-20), 4611-4622.
- Sandoval Serrano, L. J. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica; no. 11*.
- Scobie, D. R., Grosvenor, A. J., Bray, A. R., Tandon, S. K., Meade, W. J., & Cooper, A. M. B. (2015). A review of wool fibre variation across the body of sheep and the effects on wool processing. *Small Ruminant Research*, 133, 43-53.
- Shahinfar, S., & Kahn, L. (2018). Machine learning approaches for early prediction of adult wool growth and quality in Australian Merino sheep. *Computers and electronics in agriculture*, 148, 72-81.



- Shahinfar, S., & Kahn, L. (2018). Machine learning approaches for early prediction of adult wool growth and quality in Australian Merino sheep. *Computers and electronics in agriculture*, 148, 72-81.
- Shakyawar, D. B., Kadam, V., Surya, A. K., Ahmed, A., Pareek, P. K., & Temani, P. (2013). Precise measurement of wool fibre diameter using computerized projection microscope. *Indian Journal of Small Ruminants* 679 (The), 19(2), 190-192.
- Sienra, I., Neimaur, K., Kremer, R., & Urioste, J. I. (2011). Medullated fibres and fleece characteristics in Corriedale hoggets from two flocks in Uruguay. *Animal production science*, 51(11), 1034-1038.
- Valbonesi, A., Cristofanelli, S., Pierdominici, F., Gonzales, M., & Antonini, M. (2010). Comparison of fiber and cuticular attributes of alpaca and llama fleeces. *Textile Research Journal*, 80(4), 344-353.
- Valdivia, R., Álvarez, Y., & Justiniani, M. (9 de mayo de 2024). Medulación de la Fibra de Alpaca. Recuperado el 23 de noviembre de 2024, de Centro de Innovación Productiva y Transferencia Tecnológica Textil Camélidos Cusco: <https://www.gob.pe/institucion/citetextilcamelidos-cusco/noticias/951522-la-medulacion-de-la-fibra-de-alpaca>
- Vásquez, R., Gómez-Quispe, O. E., & Quispe, E. (2015). Características Tecnológicas de la Fibra Blanca de Alpaca Huacaya en la Zona Altoandina de Apurímac: Technological Characteristics of the White Fibre of Huacaya Alpaca in Theandean Region of Apurimac. *Revista de Investigaciones Veterinarias del Perú*, 26(2), 213-220.
- Vega, M. Á., Mora, L. M. Q., & Badilla, M. V. C. (2020). Inteligencia artificial y aprendizaje automático en medicina. *Revista médica sinergia*, 5(8), e557-e557.
- Villaruel, J. 1963, Un estudio de la fibra de alpaca. *Anales Científicos de la Universidad Nacional Agraria La Molina – UNALM* 1(3):246-274.
- Wang, H., Liu, X., & Wang, X. (2005). Internal structure and pigment granules in colored alpaca fibers. *Fibers and polymers*, 6, 263-268.



- Wang, L., Singh, A., & Wang, X. (2008). Dehairing Australian alpaca fibres with a cashmere dehairing machine. *Journal of the Textile Institute*, 99(6), 539-544.
- Xing, W., Liu, Y., Xin, B., Zang, L., & Deng, N. (2022). The application of deep and transfer learning for identifying cashmere and wool fibers. *Journal of Natural Fibers*, 19(1), 88-104.
- Yildiz, K. (2020). Identification of wool and mohair fibres with texture feature extraction and deep learning. *IET Image Processing*, 14(2), 348-353.



## ANEXOS

### ANEXO 1. Parámetros utilizados en el software Orange Data Mining

File	Anexo A: Reporte de Análisis en Orange Data Mining	Sat Sep 07 24, 07:56:48
<b>File</b>		
File name: C:/Users/usuario/Downloads/MEDULACION_PREDICCION.xlsx Format: Microsoft Excel spreadsheet		
<b>Data</b>		
Data instances: 439 Features: DMF, MED, DED, CVD, Fenotipo Meta attributes: ID Muestra		
<i>Data inicial ingresada al programa Orange Data Mining.</i>		



Data Table

Sat Sep 07 24, 07:58:29

Data instances: 439  
Features: 4  
Meta attributes: 1  
Target: Numeric variable 'MED'

	MED	ID Muestra	DMF	DED	CVD	Fenotipo
1	12.6697	844	13.88	3.50	25.2161	Alta
2	8.0292	787	14.22	3.75	26.3713	Media
3	13.1797	862	14.84	4.14	27.8976	Alta
4	5.51181	997	14.89	3.58	24.043	Baja
5	8.37209	1002	14.89	3.76	25.2518	Media
6	13.1609	698	14.95	3.55	23.7458	Alta
7	12.963	806	15.02	4.20	27.9627	Alta
8	12.61	853	15.28	4.28	28.0105	Alta
9	9.5	873	15.33	4.05	26.4188	Media
10	16.5775	1035	15.35	4.69	30.5537	Alta
11	7.27273	8	15.58	4.11	26.38	Media
12	7.85645	929	15.59	3.54	22.7069	Media
13	13.2216	985	15.63	4.12	26.3596	Alta
14	13.4937	1052	15.64	4.09	26.1509	Alta
15	8.125	1007	15.67	4.47	28.5258	Media
16	9.74359	1096	15.79	3.86	24.4459	Media
17	18.4127	818	15.80	4.46	28.2278	Alta
18	11.1111	783	15.84	4.34	27.399	Alta
19	13.4463	811	15.84	3.56	22.4747	Alta
20	13.2244	945	15.86	3.94	24.8424	Alta
21	12.7087	776	15.91	4.59	28.8498	Alta
22	14.506	1034	15.93	4.86	30.5085	Alta
23	10.7407	1018	15.94	5.12	32.1205	Media
24	22.6744	957	15.98	4.82	30.1627	Alta
25	11.883	1059	16.06	4.54	28.269	Media
26	4.91356	865	16.07	3.95	24.58	Baja
27	6.88226	12068	16.07	4.42	27.5047	Baja
28	12.2862	876	16.07	4.64	28.8737	Media
29	12.7148	866	16.08	4.12	25.6219	Media
30	12.4796	991	16.11	4.63	28.7399	Media
31	13.7931	1048	16.13	4.10	25.4185	Media
32	14.4781	784	16.14	5.19	32.1561	Alta
33	13.1213	782	16.16	4.60	28.4653	Alta
34	17.6912	1047	16.17	4.25	26.2832	Alta
35	13.6862	1093	16.18	4.37	27.0087	Alta
36	26.1307	922	16.20	6.05	37.3457	Alta
37	9.9481	1001	16.23	5.17	31.8546	Baja
38	18.8876	834	16.25	4.45	27.3846	Alta
39	21	798	16.29	3.94	24.1866	Alta
40	16.25	1095	16.31	4.12	25.2606	Alta
41	27.4043	708	16.31	4.37	26.7934	Alta
42	17.803	894	16.32	5.41	33.1495	Alta
43	12.1751	975	16.35	3.89	23.792	Media
44	26.4151	1043	16.38	4.34	26.4957	Alta
45	17.59	765	16.43	3.99	24.2848	Alta
46	28.3665	752	16.45	5.01	30.4559	Alta
47	9.53717	885	16.47	4.18	25.3795	Media
48	19.6571	726	16.50	4.52	27.3939	Alta
49	31.0174	994	16.51	4.24	25.6814	Alta
50	19.7201	793	16.53	4.24	25.6503	Media
51	11.7586	1019	16.57	4.10	24.7435	Baja
52	18.1399	2703	16.57	5.01	30.2354	Media
53	10.3627	995	16.58	5.91	35.6454	Baja
54	12.9341	3718	16.58	3.93	23.7033	Baja
55	12.4197	733	16.62	4.48	26.9555	Baja
56	9.50119	821	16.63	3.87	23.2712	Baja



57	24.4824	751	16.65	5.07	30.4505	Media
58	13.8288	1030	16.68	3.82	22.9017	Baja
59	17.4566	973	16.76	6.07	36.2172	Media
60	5.01193	879	16.77	4.76	28.384	Baja
61	21.8814	839	16.77	4.43	26.4162	Media
62	31.1491	720	16.78	4.62	27.5328	Media
63	9.89011	1097	16.83	4.34	25.7873	Baja
64	17.8205	12091	16.83	3.73	22.1628	Media
65	19.4614	1017	16.83	3.96	23.5294	Media
66	16.5441	1057	16.84	5.11	30.3444	Baja
67	16.4997	890	16.90	4.05	23.9645	Baja
68	21.1966	1040	16.90	4.96	29.3491	Media
69	31.4908	1094	16.92	4.44	26.2411	Alta
70	15.4812	789	16.93	4.78	28.2339	Baja
71	22.8464	1055	16.94	4.68	27.6269	Media
72	16.8453	896	16.95	4.80	28.3186	Media
73	10.168	1097	16.96	4.25	25.059	Baja
74	22.4299	824	16.96	5.41	31.8986	Media
75	24.94	999	16.96	4.48	26.4151	Media
76	14.3775	816	16.97	4.21	24.8085	Baja
77	26.2201	1079	16.97	4.07	23.9835	Media
78	11.0638	892	16.99	6.08	35.7858	Baja
79	12.395	846	16.99	4.17	24.5438	Baja
80	22.973	770	17.00	4.50	26.4706	Media
81	31.068	852	17.05	5.53	32.434	Alta
82	15.2276	951	17.06	4.45	26.0844	Baja
83	23.38	927	17.09	4.28	25.0439	Media
84	16.2602	914	17.11	4.72	27.5862	Baja
85	30.9423	802	17.14	5.04	29.4049	Alta
86	20.1842	743	17.15	4.39	25.5977	Media
87	18.9672	1072	17.16	4.40	25.641	Media
88	23.5165	987	17.17	4.54	26.4415	Media
89	12.3703	823	17.19	4.50	26.178	Baja
90	25.9768	2705	17.21	4.57	26.5543	Media
91	18.2482	753	17.27	4.23	24.4933	Media
92	9.86842	844	17.28	3.39	19.6181	Baja
93	25	724	17.28	4.26	24.6528	Media
94	18.1366	1023	17.29	5.04	29.1498	Media
95	29.8477	974	17.31	5.21	30.0982	Media
96	37.1429	258	17.32	5.33	30.7737	Alta
97	40.3226	921	17.32	5.17	29.8499	Alta
98	11.8325	1087	17.33	4.02	23.1968	Baja
99	15.9134	757	17.33	4.61	26.6013	Baja
100	26.3374	988	17.33	4.70	27.1206	Media

+ 339 more

Tabla de datos inicial, 3 variables predictoras, 2 variable objetivo, 1 variable categórica y 1 meta atributo.

Preproces

Sat Sep 07 24, 08:02:58

Settings

Impute Missing Values: Average/Most frequent

Se imputaron valores faltantes con los promedios, y se eliminaron datos atípicos.



Feature Statistics

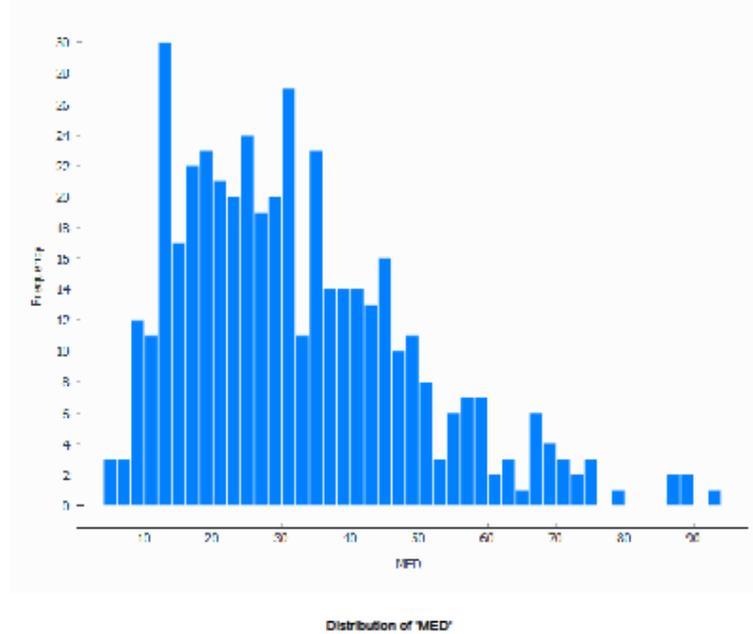
Sat Sep 07 24, 08:04:22

Name	Distribution	Mean	Mode	Median	Dispersion	Min.	Max.	Missing
DMF		18.8377	17.48	18.62	0.1072	13.88	29.17	0 (0%)
DED		5.1170	4.34	5.00	0.1797	3.39	12.25	0 (0%)
CVD		27.1664	26.4957	26.7276	0.136061	17.2051	43.6523	0 (0%)
MED		32.1667	11.1111	29.5675	0.524615	4.91356	92.8789	0 (0%)
Fenotipo			Alta		1			0 (0%)

Análisis exploratorio de las estadísticas descriptivas de los datos.

Distributions

Sat Sep 07 24, 08:05:36

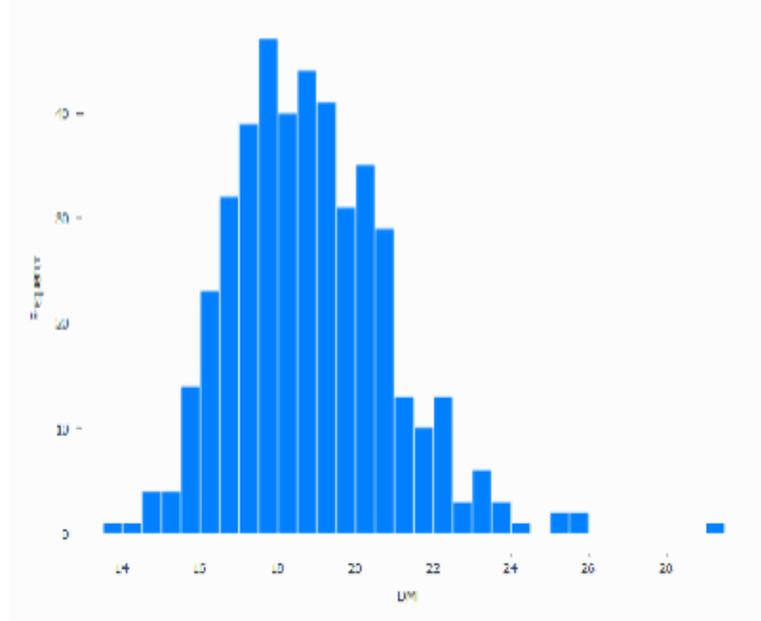


Distribución de los datos de Medulación.



Distributions

Sat Sep 07 24, 08:06:36

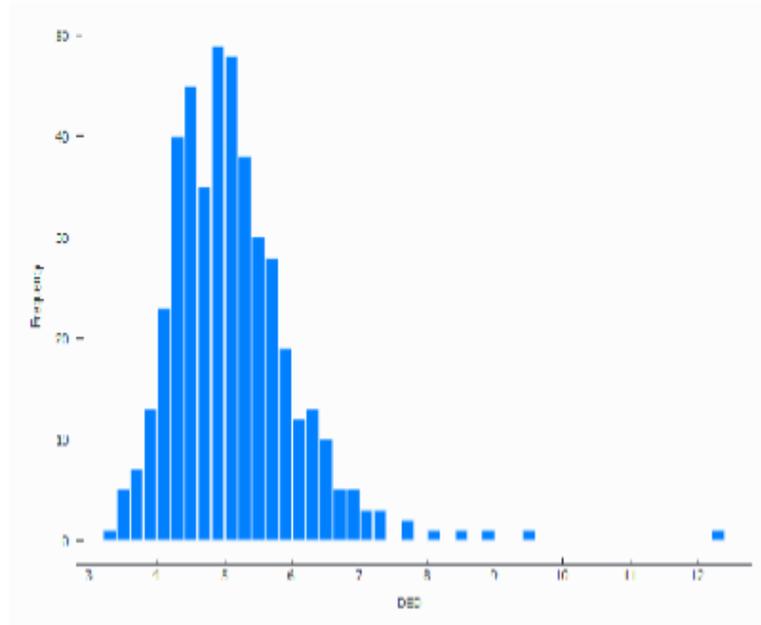


Distribution of 'DMP'

Distribución de los datos de finura de fibra.

Distributions

Sat Sep 07 24, 08:07:25



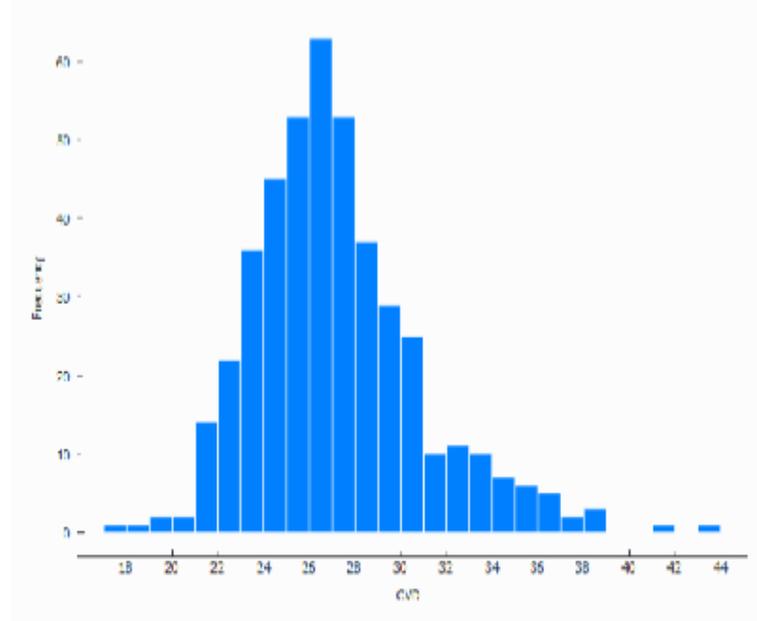
Distribution of 'DED'



Distribución de los datos de desviación estándar del diámetro.

Distributions

Sat Sep 07 24, 08:08:28



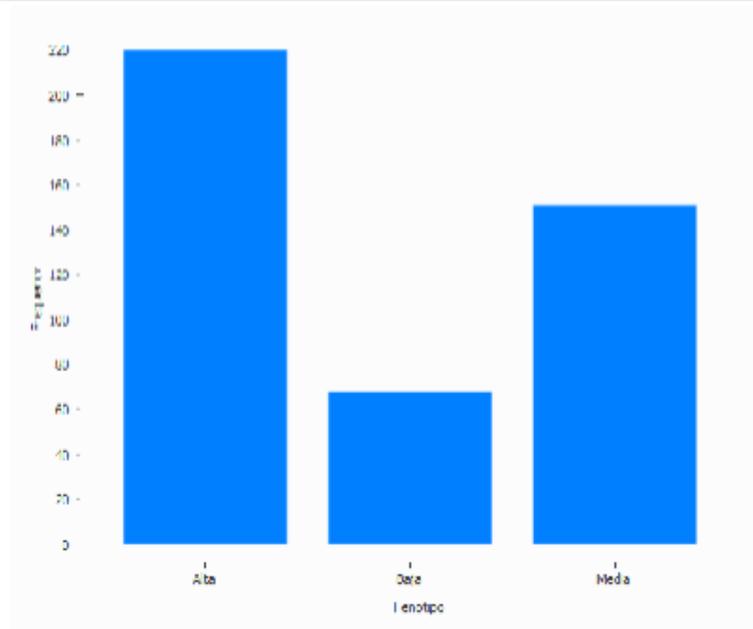
Distribution of 'CVD'

Distribución de los datos en el coeficiente de variación del diámetro.



Distributions

Sat Sep 07 24, 08:09:40



Distribution of 'Fenotipo'

Frecuencia encontradas en la variable categórica "Fenotipo".

Gradient Boosting

Sat Sep 07 24, 08:10:37

Name: Gradient Boosting

Model parameters

Method: Gradient Boosting (scikit-learn)  
Number of trees: 105  
Learning rate: 0.1  
Replicable training: Yes  
Maximum tree depth: 3  
Fraction of training instances: 1  
Stop splitting nodes with maximum instances: 2

Data

Data instances: 439  
Features: DMF, DED, CVD, Fenotipo  
Meta attributes: ID Muestra  
Target: MED

Parámetros utilizados en el modelo.



**kNN** Sat Sep 07 24, 08:11:21

Name: kNN

**Model parameters**

---

Number of neighbours: 20  
Metric: Euclidean  
Weight: By Distances

**Data**

---

Data instances: 439  
Features: DMF, DED, CVD, Fenotipo  
Meta attributes: ID Muestra  
Target: MED

Parámetros utilizados en el modelo.

**Random Forset** Sat Sep 07 24, 08:12:10

Name: Random Forest

**Model parameters**

---

Number of trees: 50  
Maximal number of considered features: unlimited  
Replicable training: No  
Maximal tree depth: unlimited  
Stop splitting nodes with maximum instances: 5

**Data**

---

Data instances: 439  
Features: DMF, DED, CVD, Fenotipo  
Meta attributes: ID Muestra  
Target: MED

Parámetros utilizados en el modelo.

**AdaBoost** Sat Sep 07 24, 08:12:55

Name: AdaBoost

**Model parameters**

---

Base estimator: tree  
Number of estimators: 50  
Algorithm (classification): Samme.r  
Loss (regression): Square

**Data**

---

Data instances: 439  
Features: DMF, DED, CVD, Fenotipo  
Meta attributes: ID Muestra  
Target: MED

Parámetros utilizados en el modelo.



### Tree

Sat Sep 07 24, 08:13:12

Name: Tree

#### Model parameters

Pruning: at least 2 instances in leaves, at least 5 instances in internal nodes, maximum depth 100  
Splitting: Stop splitting when majority reaches 95% (classification only)  
Binary trees: Yes

#### Data

Data instances: 439  
Features: DMF, DED, CVD, Fenotipo  
Meta attributes: ID Muestra  
Target: MED

Parámetros utilizados en el modelo.

### Neural Network

Sat Sep 07 24, 08:13:41

Name: Neural Network

#### Model parameters

Hidden layers: 120  
Activation: ReLu  
Solver: Adam  
Alpha: 0.0001  
Max iterations: 200  
Replicable training: Yes

#### Data

Data instances: 439  
Features: DMF, DED, CVD, Fenotipo  
Meta attributes: ID Muestra  
Target: MED

Parámetros utilizados en el modelo.

### Linear Regression

Sat Sep 07 24, 08:13:57

Name: Linear Regression

#### Model parameters

Regularization: No Regularization  
Fit intercept: Yes

#### Data

Data instances: 439  
Features: DMF, DED, CVD, Fenotipo  
Meta attributes: ID Muestra  
Target: MED

Parámetros utilizados en el modelo.



Test and Score

Sat Sep 07 24, 08:14:28

Settings

Sampling type: 3-fold Cross validation

Scores

Model	Train	Test	RMSE	MAE	R2
KNN	0.025	0.026	8.986	6.704	0.716
Gradient Boosting	0.426	0.015	6.828	4.966	0.836
Random Forest	0.348	0.024	6.515	4.739	0.851
Linear Regression	0.026	0.014	7.346	5.666	0.810
AdaBoost	0.434	0.020	7.501	5.250	0.802
Neural Network	1.615	0.013	9.085	6.869	0.710
Tree	0.373	0.000	7.809	5.710	0.786

Métricas de desempeño para 7 modelos de aprendizaje automático entrenados para la predicción de la medulación.



Data Table

Sat Sep 07 24, 08:16:24

Data Instances: 439  
Features: 4  
Meta attributes: 9  
Target: Numeric variable 'MED'

	MED	ID Muestra	KNN	Gradient Boosting	Random Forest	Linear Regression	AdaBoost	Neural Network	Tree	Fold	DMF	DED	CVD	Fenotipo
1	12.6697	844	15.6035	12.7923	13.1441	11.2879	13.4937	16.9802	13.4672	1	13.88	3.50	25.2161	Aita
2	5.51181	997	14.1226	4.65311	10.7673	-3.15838	9.86842	16.3628	12.3708	1	14.89	3.58	24.043	Baja
3	13.1609	698	15.1474	13.228	13.1441	16.5546	13.4937	21.5949	13.4672	1	14.95	3.55	23.7458	Aita
4	12.61	853	14.6032	13.7884	13.4009	20.4233	13.1213	15.2627	12.9932	1	15.28	4.28	28.0105	Aita
5	16.5775	1035	20.8499	14.2696	14.3017	21.8993	14.506	13.0914	18.5902	1	15.35	4.69	30.5537	Aita
6	9.74359	1096	16.5624	10.6066	8.64929	10.3204	8.125	20.3509	12.5005	1	15.79	3.86	24.4459	Media
7	18.4127	818	16.897	14.1668	13.4891	23.1672	13.1213	16.8697	12.9932	1	15.80	4.46	28.2278	Aita
8	11.1111	783	18.2756	14.0033	13.6032	23.0788	13.6862	18.0308	12.9932	1	15.84	4.34	27.399	Aita
9	13.4463	811	18.0681	12.9285	13.2529	21.3449	13.6862	26.7886	13.4672	1	15.84	3.56	22.4747	Aita
10	13.2244	945	19.4736	14.6971	13.3586	22.2924	13.6862	22.0748	13.4672	1	15.86	3.94	24.8424	Aita
11	10.7407	1018	23.8056	11.3989	12.001	13.7503	14.506	18.3563	12.5005	1	15.94	5.12	32.1205	Media
12	11.883	1059	16.4305	13.3893	11.7047	13.0724	13.1213	18.8541	12.5005	1	16.06	4.54	28.269	Media
13	4.91356	865	15.8831	9.40927	11.2234	3.89088	10.3627	14.3657	12.3708	1	16.07	3.95	24.58	Baja
14	6.88226	12068	15.8433	9.60197	11.3109	4.80107	12.2862	12.4511	12.3708	1	16.07	4.42	27.5047	Baja
15	12.4796	991	15.4932	11.7521	12.6531	13.4696	12.7148	18.7764	18.4096	1	16.11	4.63	28.7399	Media
16	13.7931	1048	17.3647	10.6629	12.3274	12.5713	12.7148	19.8628	19.5832	1	16.13	4.10	25.4185	Media
17	14.4781	784	27.0642	18.332	18.2065	26.0706	17.803	16.6149	18.5902	1	16.14	5.19	32.1561	Aita
18	17.6912	1047	21.2142	14.5193	16.4017	24.4791	17.59	20.866	13.4672	1	16.17	4.25	26.2832	Aita
19	16.25	1095	22.4098	20.3179	22.8715	24.9652	17.803	23.0058	24.9398	1	16.31	4.12	25.2606	Aita
20	12.1751	975	18.393	16.9027	17.6992	13.3755	18.9672	20.7767	19.5832	1	16.35	3.89	23.792	Media
21	28.3665	752	26.3258	21.6251	20.9451	27.0171	19.6571	18.4189	18.6236	1	16.45	5.01	30.4559	Aita
22	9.53717	885	19.5024	17.8427	18.2101	14.4618	19.4614	19.9	19.5832	1	16.47	4.18	25.3795	Media
23	9.50119	821	16.9654	10.3202	11.4964	6.82657	10.3627	14.7541	7.45102	1	16.63	3.87	23.2712	Baja
24	24.4824	751	23.5638	17.3938	18.2776	16.5237	18.1399	18.4878	18.4096	1	16.65	5.07	30.4505	Media
25	13.8288	1030	17.3559	11.6782	10.975	7.05016	9.89011	15.0732	7.45102	1	16.68	3.82	22.9017	Baja
26	21.8814	839	23.834	19.2782	19.5384	16.3337	19.7201	19.5057	28.0446	1	16.77	4.43	26.4162	Media
27	17.8205	12091	19.8696	21.3741	20.4163	15.9096	18.9693	22.7175	19.5832	1	16.83	3.73	22.1628	Media
28	16.4997	890	17.0704	14.1864	14.307	8.55177	11.8325	13.7822	15.9281	1	16.90	4.05	23.9645	Baja
29	22.8464	1055	25.9946	23.736	22.2802	17.4428	28.3144	19.1171	18.4096	1	16.94	4.68	27.6269	Media
30	10.168	1097	16.414	14.9192	14.5319	9.06988	12.4197	12.7594	15.9281	1	16.96	4.25	25.059	Baja
31	22.4299	824	25.4372	21.6087	21.5696	18.1998	20.1397	18.3581	18.4096	1	16.96	5.41	31.8986	Media
32	14.3775	816	15.7417	15.4156	14.5319	9.08854	12.4197	12.9616	15.9281	1	16.97	4.21	24.8085	Baja
33	26.2201	1079	20.9624	21.8703	21.0135	17.0481	19.8386	20.7627	19.5832	1	16.97	4.07	23.9835	Media
34	11.0638	892	24.362	13.3086	15.0024	10.8222	13.6646	9.26777	16.724	1	16.99	6.08	35.7858	Baja
35	12.395	846	16.1677	15.7456	14.5854	9.16342	11.1111	13.1625	15.9281	1	16.99	4.17	24.5438	Baja
36	22.973	770	24.5401	24.2071	22.7815	17.5859	23.5165	19.5462	28.0446	1	17.00	4.50	26.4706	Media
37	23.38	927	21.8802	22.7783	21.2635	17.8869	22.807	20.1202	19.5832	1	17.09	4.28	25.0439	Media
38	30.9423	802	28.1113	32.4054	32.6162	30.1272	31.0811	22.7431	30.9866	1	17.14	5.04	29.4049	Aita
39	18.2482	753	23.2612	24.744	24.3097	18.8508	20.7641	20.4807	22.857	1	17.27	4.23	24.4933	Media
40	18.1366	1023	22.9546	21.3892	23.4764	19.4044	22.082	18.7592	21.5078	1	17.29	5.04	29.1498	Media
41	29.8477	974	26.5615	25.4101	23.937	19.5889	23.0961	18.6391	21.5078	1	17.31	5.21	30.0982	Media
42	8.47107	697	16.4624	15.1915	13.9625	11.3257	12.8749	12.264	14.4063	1	17.36	4.34	25	Baja
43	21.6667	RA-56	23.9954	25.174	24.0604	19.7272	23.5165	20.3832	22.857	1	17.42	4.30	24.6843	Media
44	21.6667	1066	29.6838	24.9613	23.8324	20.8699	20.7641	22.2193	21.3251	1	17.47	6.72	38.4659	Media
45	22.2222	803	25.461	24.8181	25.9622	20.2045	23.5165	19.2007	27.5668	1	17.47	4.85	27.7619	Media
46	22.963	968	22.7659	23.4066	23.8821	19.9518	23.5165	21.238	22.857	1	17.47	4.14	23.6978	Media
47	29.2007	701	27.3894	23.9071	25.3384	20.459	21.5385	18.559	28.5289	1	17.48	5.44	31.1213	Media
48	46.3511	860	27.3512	36.3923	34.8052	31.6788	31.4908	25.6094	32.122	1	17.48	4.83	27.6316	Aita
49	29.5675	886	24.0074	23.5706	23.6222	20.1988	23.0961	19.8777	23.0543	1	17.49	4.53	25.9005	Media
50	30.6554	875	28.8271	34.4476	36.1521	32.075	33.8832	24.3129	31.8252	1	17.49	5.86	33.5049	Aita
51	15.1261	809	16.2155	15.0927	15.494	12.2109	14.9254	10.8725	17.7341	1	17.50	4.66	26.6286	Baja
52	34.0206	797	29.4677	38.3358	37.2486	31.988	37.8058	27.0279	35.7058	1	17.55	4.65	26.4957	Aita
53	13.7342	887	17.3415	15.9175	15.791	12.7058	14.9254	9.79096	13.6994	1	17.58	4.95	28.157	Baja
54	14.1643	739	18.362	15.0136	14.9762	13.0786	12.8749	14.2057	13.9379	1	17.68	4.00	22.6244	Baja
55	18.9364	840	16.3209	15.7195	15.163	13.1348	13.6646	11.9265	14.4063	1	17.68	4.40	24.8869	Baja



56	16.8498	1062	17.6852	15.3686	15.048	13.5715	12.8749	13.4656	13.9379	1	17.76	4.12	23.1982	Baja
57	41.115	833	25.6437	36.6158	36.1192	33.1494	31.7593	32.4555	33.4089	1	17.77	4.19	23.5791	Alta
58	29.0398	775	26.3548	23.6553	24.4229	21.8581	25.2568	19.4968	21.3251	1	17.78	6.22	34.9831	Media
59	29.2155	1022	24.9529	23.3148	25.0252	22.217	25.8824	19.1634	21.5078	1	17.86	5.06	28.3315	Media
60	27.3824	982	25.5934	25.3921	25.045	22.2789	21.5385	19.5574	21.6344	1	17.87	4.84	27.0845	Media
61	13.8211	800	23.5062	17.2114	13.5389	14.3206	13.6646	17.6874	13.9379	1	17.88	3.63	20.302	Baja
62	20.0234	1004	24.3411	25.5057	24.8436	22.3476	23.0961	20.0989	26.3758	1	17.88	4.58	25.6152	Media
63	23.3221	905	26.0609	29.4365	25.3829	22.3556	25.2568	18.8832	21.5078	1	17.89	5.26	29.4019	Media
64	15.2434	851	18.516	15.805	15.002	14.41	12.8749	13.3717	13.9379	1	17.90	4.14	23.1285	Baja
65	30.8296	899	25.6588	25.1479	27.5437	23.3713	25.8824	19.5121	31.4912	1	18.08	4.97	27.4889	Media
66	20.5827	947	22.2758	24.2822	25.8312	23.8636	25.3266	23.9304	26.0962	1	18.12	4.02	22.1854	Media
67	24.5283	841	23.9749	24.1913	25.9288	24.379	25.3266	22.0232	26.0962	1	18.22	4.30	23.6004	Media
68	31.9797	877	35.5756	36.5693	33.0244	35.692	30.5913	30.357	30.9866	1	18.25	5.01	27.4521	Alta
69	35.6775	835	40.495	39.8593	40.7501	35.5927	43.6214	29.4532	47.5427	1	18.27	5.48	29.9945	Alta
70	29.0984	226	25.0324	23.8578	25.3504	24.7546	25	21.5469	26.3758	1	18.29	4.40	24.0569	Media
71	18.4265	946	22.8872	15.4376	13.7585	16.6945	14.1431	8.68668	11.5376	1	18.36	5.15	28.0501	Baja
72	13.5321	710	20.341	14.8956	13.7706	17.2483	12.8749	13.2562	13.9379	1	18.37	4.21	22.9178	Baja
73	49.6423	981	35.3185	37.0251	34.4445	36.8059	34.8281	32.1749	36.5073	1	18.44	4.86	26.3557	Alta
74	38.1951	883	41.4126	40.3412	43.962	36.7052	43.6214	31.0764	40.6769	1	18.48	5.37	29.0584	Alta
75	46.6667	757	37.7166	39.1484	38.1798	37.269	37.8058	32.6182	36.5073	1	18.53	4.92	26.5515	Alta
76	49.8958	967	42.1335	41.7888	43.1886	36.8614	48.4202	31.0458	47.5427	1	18.56	5.73	30.8728	Alta
77	29.3737	1060	29.0712	23.4201	23.7381	25.259	24.6622	20.2	21.3251	1	18.60	6.22	33.4409	Media
78	30.6818	1077	23.949	25.0231	26.9501	26.3285	24.3605	19.9008	25.2334	1	18.64	5.08	27.2532	Media
79	41.7085	1048	34.7936	37.2466	40.771	37.3229	39.2727	31.6887	39.2822	1	18.68	5.88	31.4775	Alta
80	19.4556	908	25.402	25.3598	25.1271	27.6125	25.3266	24.0704	26.0962	1	18.74	4.24	22.6254	Media
81	25.0648	829	30.6132	25.6673	24.808	25.327	24.6622	24.2343	21.3251	1	18.75	6.90	36.8	Media
82	41.3146	830	34.0959	36.3414	38.9918	38.0352	39.2727	32.7294	31.7807	1	18.84	5.94	31.5287	Alta
83	40.0731	984	33.8713	37.971	38.5476	40.1497	31.7593	38.7847	49.9132	1	18.94	4.39	23.1785	Alta
84	40.9821	729	40.8912	41.3681	44.1167	39.9917	49.9093	35.4767	49.2715	1	19.08	5.27	27.6205	Alta
85	37.0166	1037	36.83	37.8264	35.5226	40.9377	31.4659	36.8468	39.1369	1	19.22	5.14	26.743	Alta
86	33.8188	906	42.3676	38.5788	38.0533	40.6425	35.1449	36.1705	33.764	1	19.23	5.41	28.1331	Alta
87	34.1853	1038	41.4407	46.3817	42.8889	40.8358	45.6161	36.3917	54.5579	1	19.28	5.47	28.3714	Alta
88	48.3112	746	38.3396	42.8512	41.8242	41.3678	38.1381	37.36	54.5579	1	19.30	5.16	26.7358	Alta
89	44.1264	964	39.3545	40.3318	38.8407	41.992	38.1381	38.5617	36.5073	1	19.35	4.92	25.4264	Alta
90	25.4717	RA-54	26.5426	27.6134	26.3206	30.3095	27.9368	20.9392	28.1763	1	19.36	5.14	26.5496	Media
91	35.141	1061	40.0271	41.0638	40.8856	40.746	38.7931	36.4791	38.4361	1	19.39	5.95	30.6859	Alta
92	33.0109	925	32.3005	35.0863	34.199	40.0226	37.5519	36.5854	31.7807	1	19.44	6.61	34.0021	Alta
93	41.0588	998	42.1847	40.9001	43.9104	41.2831	39.3204	36.9975	38.4361	1	19.44	5.76	29.6296	Alta
94	43.1006	889	41.7499	43.4233	45.8138	42.2679	46.7662	38.3846	54.5579	1	19.48	5.25	26.9507	Alta
95	25.6959	1069	29.3704	21.9344	24.5556	29.1134	23.7721	22.6966	25.3226	1	19.51	6.48	33.2137	Media
96	44.74	1068	41.9369	41.1284	43.1272	43.0255	46.8421	39.8916	36.5073	1	19.51	4.87	24.9616	Alta
97	41.1765	955	42.6118	41.8386	40.3516	42.2918	42.6829	38.2429	54.5579	1	19.55	5.49	28.0818	Alta
98	29.9712	1083	40.76	39.0066	37.8967	43.3406	38.1381	40.0006	30.9866	1	19.61	5.05	25.7522	Alta
99	20.4117	920	29.5302	25.0288	23.904	32.5294	23.7607	23.0862	28.1763	1	19.65	4.82	24.5293	Media
100	33.3333	909	33.1693	33.0215	33.1655	41.3505	33.8832	38.2071	31.7807	1	19.69	6.51	33.0625	Alta

+ 339 more

Valores reales frente a valores predichos por cada modelo de aprendizaje automático.

Save Data

Sat Sep 07 24, 08:18:25

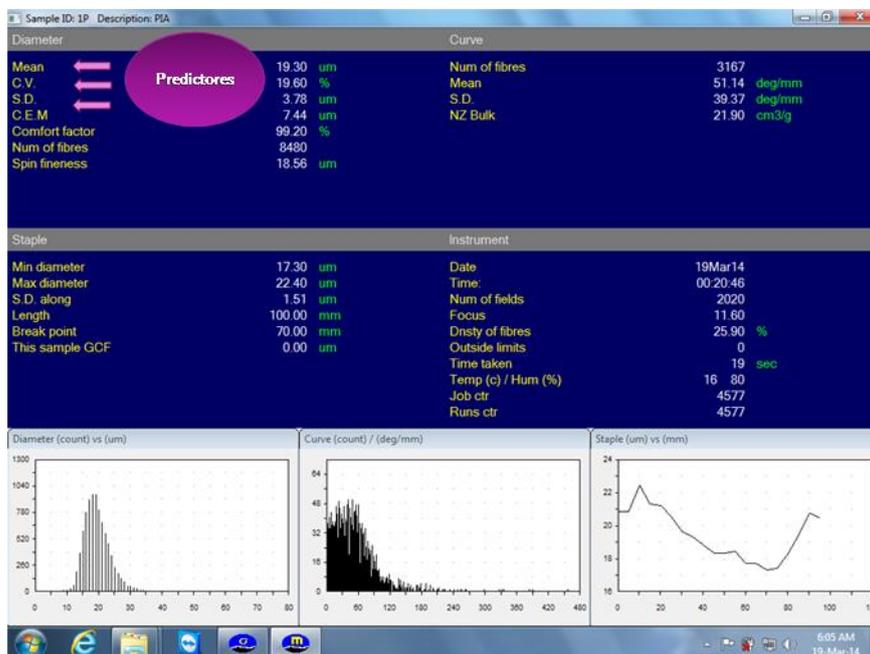
Data instances: 439  
Features: 4  
Meta attributes: 9  
Target: Numeric variable 'MED'

File name: C:/Users/usuario/Downloads/Resultado\_Predicciones\_Medulación.csv  
Format: Comma-separated values  
Type annotations: Yes

## ANEXO 2. Panel fotográfico del uso y procesamiento de OFDA 2000 y FIBER MED



Lectura de fibra de alpaca en OFDA2000



Interfaz gráfica de OFDA2000



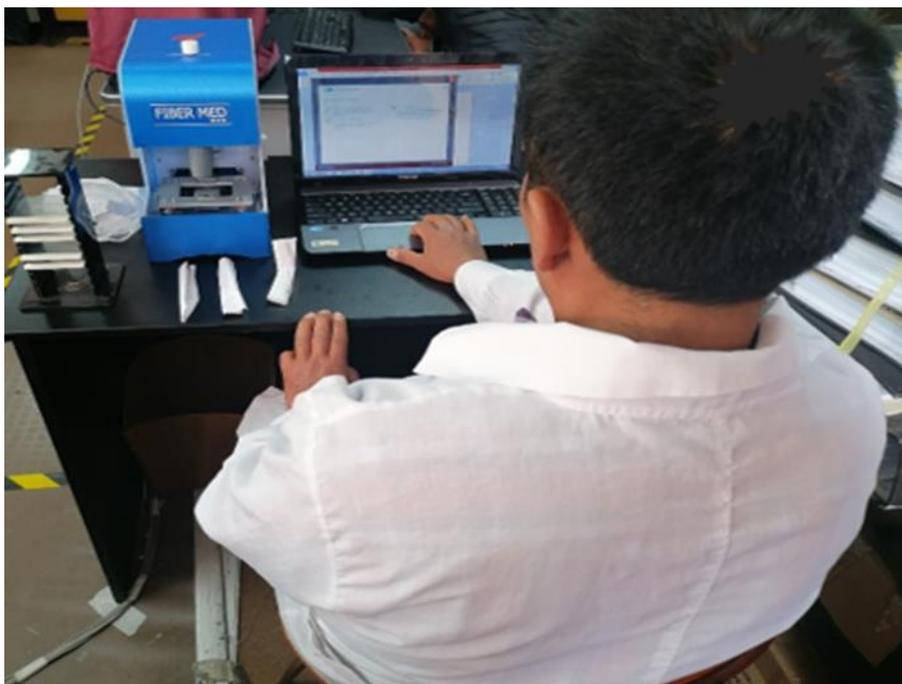
Preparación de Slides y preparación de lectura en OFDA2000



Posición vertical de las muestras de fibra previo a lectura.



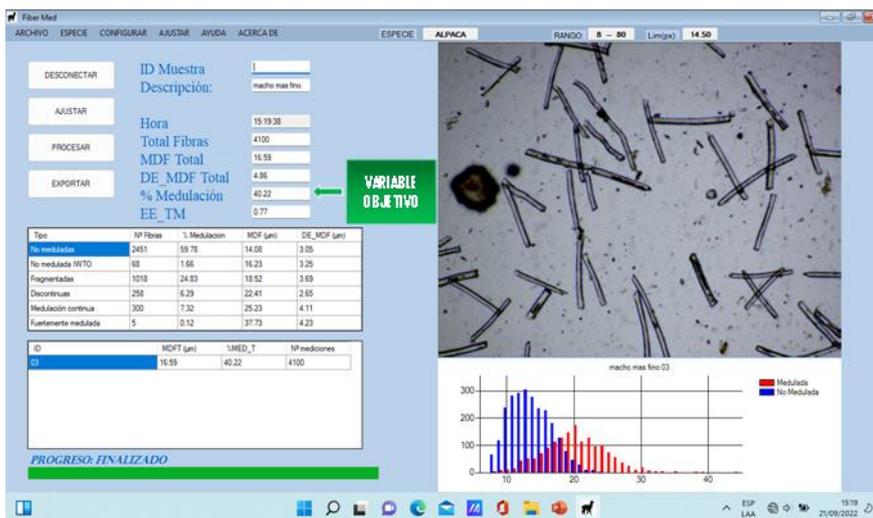
Vista posterior del FIBERMED



Vista Frontal del FIBERMED, en plena lectura de muestra.



Preparación de muestras para lectura en FIBERMED.



Interfaz gráfica del medulador automático FIBERMED.



### ANEXO 3. Declaración jurada de autenticidad de tesis



Universidad Nacional  
del Altiplano Perú



Vicerrectoría  
de Investigación



Repositorio  
Institucional

#### DECLARACIÓN JURADA DE AUTENTICIDAD DE TESIS

Por el presente documento, Yo CARLOS ALBERTO VARGAS MAMANI  
identificado con DNI: 89028450 en mi condición de egresado de:

Escuela Profesional,  Programa de Segunda Especialidad,  Programa de Maestría o Doctorado  
MEDICINA VETERINARIA Y ZOOTECNIA

informo que he elaborado esta  Tesis o  Trabajo de Investigación denominada:  
" ESTIMACIÓN DE LA MODULACIÓN EN FIBRA DE ALPACA HUACAYA  
BLANCO, MEDIANTE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO.

Es un tema original.

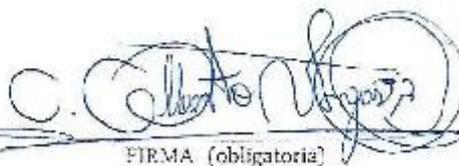
Declaro que el presente trabajo de tesis es elaborado por mí persona y no existe **plagio/copia** de ninguna naturaleza, en especial de otro documento de investigación (tesis, revista, texto, congreso, o similar) presentado por persona natural o jurídica alguna ante instituciones académicas, profesionales, de investigación o similares, en el país o en el extranjero.

Dejo constancia que las citas de otros autores han sido debidamente identificadas en el trabajo de investigación, por lo que no asumiré como propias las opiniones vertidas por terceros, ya sea de fuentes encontradas en medios escritos, digitales o Internet.

Asimismo, ratifico que soy plenamente consciente de todo el contenido de la tesis y asumo la responsabilidad de cualquier error u omisión en el documento, así como de las consecuencias éticas y legales involucradas.

En caso de incumplimiento de esta declaración, me someto a las disposiciones legales vigentes y a las sanciones correspondientes de igual forma me someto a las sanciones establecidas en las Directivas y otras normas internas, así como las que me alcancen del Código Civil y Normas Legales conexas por el incumplimiento del presente compromiso

Perú, 05 de NOVIEMBRE del 2024

  
FIRMA (obligatoria)



Huella



## ANEXO 4. Autorización para el depósito de tesis en el Repositorio Institucional



Universidad Nacional  
del Altiplano Puno



Vicerrectorado  
de Investigación



Repositorio  
Institucional

### AUTORIZACIÓN PARA EL DEPÓSITO DE TESIS O TRABAJO DE INVESTIGACIÓN EN EL REPOSITORIO INSTITUCIONAL

Por el presente documento, Yo CARLOS ALBERTO VARGAS MATAZI  
identificado con DNI 80028450 en mi condición de egresado de:

Escuela Profesional,  Programa de Segunda Especialidad,  Programa de Maestría o Doctorado  
MEDICINA VETERINARIA Y ZOOTECNIA

informo que he elaborado el/a  Tesis o  Trabajo de Investigación denominada:

ESTIMACIÓN DE LA MODULACIÓN EN FIBRA DE ALPACA HUACAYA  
BLANCO, MEDIANTE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO

para la obtención de  Grado,  Título Profesional o  Segunda Especialidad.

Por medio del presente documento, afirmo y garantizo ser el legítimo, único y exclusivo titular de todos los derechos de propiedad intelectual sobre los documentos arriba mencionados, las obras, los contenidos, los productos y/o las creaciones en general (en adelante, los "Contenidos") que serán incluidos en el repositorio institucional de la Universidad Nacional del Altiplano de Puno.

También, doy seguridad de que los contenidos entregados se encuentran libres de toda contraseña, restricción o medida tecnológica de protección, con la finalidad de permitir que se puedan leer, descargar, reproducir, distribuir, imprimir, buscar y enlazar los textos completos, sin limitación alguna.

Autorizo a la Universidad Nacional del Altiplano de Puno a publicar los Contenidos en el Repositorio Institucional y, en consecuencia, en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, sobre la base de lo establecido en la Ley N° 30035, sus normas reglamentarias, modificatorias, sustitutorias y conexas, y de acuerdo con las políticas de acceso abierto que la Universidad aplique en relación con sus Repositorios Institucionales. Autorizo expresamente toda consulta y uso de los Contenidos, por parte de cualquier persona, por el tiempo de duración de los derechos patrimoniales de autor y derechos conexos, a título gratuito y a nivel mundial.

En consecuencia, la Universidad tendrá la posibilidad de divulgar y difundir los Contenidos, de manera total o parcial, sin limitación alguna y sin derecho a pago de contraprestación, remuneración ni regalía alguna a favor mío; en los medios, canales y plataformas que la Universidad y/o el Estado de la República del Perú determine, a nivel mundial, sin restricción geográfica alguna y de manera indefinida, pudiendo crear y/o extraer los metadatos sobre los Contenidos, e incluir los Contenidos en los índices y buscadores que estimen necesarios para promover su difusión.

Autorizo que los Contenidos sean puestos a disposición del público a través de la siguiente licencia:

Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visitar: <https://creativecommons.org/licenses/by-nc-sa/4.0/>

En señal de conformidad, suscribo el presente documento.

Puno, 05 de NOVIEMBRE del 2024

C. Alberto Vargas  
FIRMA (obligatoria)



Huella