

Universidad de Lima  
Facultad de Ingeniería  
Carrera de Ingeniería Industrial



# **MACHINE LEARNING APPLIED TO MILK SAMPLE CLASSIFICATION**

Tesis para optar el Título Profesional de Ingeniero Industrial

**Mia Leonarda Leon Loyola**

**Código 20172282**

**Diego Daniel Ossa De la Cruz**

**Código 20172430**

**Asesor**

**José Antonio Taquía Gutiérrez**

Lima – Perú

Mayo de 2023

<b>Propuesta</b> <b>Carrera Ingeniería Industrial</b>
<b>Título</b> <b>MACHINE LEARNING APPLIED TO MILK SAMPLE CLASSIFICATION</b>
<b>Autor(es)</b> <b>Mia León y Diego Ossa</b> 20172282@aloe.ulima.edu.pe; 20172430@aloe.ulima.edu.pe Universidad de Lima
<b>Resumen:</b> El documento presenta los resultados de la evaluación del proceso de clasificación de muestras de leche por medio de la modelación de técnicas de machine learning. Esta investigación tuvo como objetivo discriminar la presencia o ausencia de adulterantes, lo cual permita la obtención de lácteos adecuados para el consumo humano. Asimismo, acelerar y precisar el proceso de inspección de dichas muestras. La relevancia del presente estudio se puede comprender desde el producto sometido a análisis: la leche. Este es de consumo masivo, sobre todo, en público infantil. Por lo expuesto, se considera relevante demostrar de manera eficiente que se brinda productos de calidad a la población y este documento es un aporte a la credibilidad de la integridad de productos lácteos.
<b>Palabras Clave:</b> Leche, Calidad, Machine Learning, Random Forest, K-Nearest Neighbors, Redes Neuronales
<b>Abstract:</b> The document presents the results of the evaluation of the classification process of milk samples through the modeling of machine learning techniques. The objective of this research was to discriminate the presence or absence of adulterants, which allowed obtaining adequate damages for human consumption. Also, speed up and specify the inspection process of said samples. The relevance of this study can be understood from the product under analysis: milk. This is for mass consumption, especially among children. Due to the above, it is considered relevant to efficiently demonstrate that quality products are provided to the population and this document is a contribution to the reliability of the integrity of dairy products.
<b>Keywords:</b> Milk, Quality, Machine Learning, Random Forest, K-Nearest Neighbors, Neural Networks
<b>Línea de investigación IDIC – ULIMA: Quality &amp; Reliability Engineering</b>
<b>Área y Sub-áreas de Investigación:</b> Diseño de sistemas automatizado para el control de calidad.
<b>Objetivo (s) de Desarrollo Sostenible (ODS) relacionado (s) al tema de investigación.</b> <b>ODS 9: Industria, innovación e infraestructura.</b>

## **PLANTEAMIENTO DEL PROBLEMA**

La presente investigación, busca clasificar muestras de leche, según la presencia o ausencia de adulterantes, que permitan la obtención de lácteos adecuados para el consumo humano mediante métodos de aprendizaje automatizado.

### **OBJETIVOS**

El propósito de la investigación es clasificar muestras de leche con el apoyo de técnicas de machine learning, que discriminarán la presencia o ausencia de adulterantes, lo cual permita la obtención de lácteos adecuados para el consumo humano. Asimismo, acelerar y precisar el proceso de inspección de dichas muestras. Los métodos de investigación y la recopilación de datos se establecieron para evaluar los datos técnicos recolectados y diseñar, con base en machine learning, un modelo automatizado, que agilice la examinación de la información recabada y brinde un resultado más efectivo sobre su condición de calidad óptima.

### **JUSTIFICACIÓN**

La leche forma parte de una notable base para para la reproducción de microorganismos a causa de su alto porcentaje nutricional; debido a esto, la sanidad de esta se debe evaluar constantemente. La leche cruda en su óptimo estado no debe poseer residuos ni sedimentos, no ser insípida ni tener color y olor inusuales; además, de un contenido mínimo de bacterias, la ausencia de sustancias químicas e índices normales de acidez en su composición. Distinguir la integridad de los lácteos promoverá la implementación de normas, que aseguren productos saludables (Gonzales et al.,2019, p.2.). La presente investigación, busca clasificar muestras de leche, según la presencia o ausencia de adulterantes, que permitan la obtención de lácteos adecuados para el consumo humano. El método tradicional para identificar propiedades de este producto se da por medio de la experimentación; sin embargo, métodos de clasificación y técnicas vinculadas al empleo de machine learning muestran mejores resultados que métodos tradicionales; además, son de gran utilidad, en la actualidad, para la sociedad, pues están transformando procesos industriales, que traen consigo mejoras en lo que refiere a operaciones productivas eficientes y toma de decisiones ágiles. Además, permite la recopilación de datos, predicción de sucesos en base a estos, creación de modelos de valor agregado por medio del empleo de técnicas de automatización, acelerar y precisar el proceso de inspección de muestras.

### **DISEÑO METODOLÓGICO**

En el presente estudio se realizó la clasificación de la leche- empleando machine learning- para la determinación de la presencia de adulterantes dentro de las muestras seleccionadas. Para el logro de este objetivo, se utilizarán métodos de clasificación como random forest, redes neuronales y k-nearest neighbor, generalmente usados por la industria de alimentos.

## NOTAS (AGRADECIMIENTOS)

A mi amada madre, Liliana. Agradecer su sacrificio y entrega durante todos estos años. Por ser mi soporte y motivación constante para no rendirme. Por ser mi inspiración, creer en mí y ayudarme a escalar cada vez más alto trazándome nuevos retos; porque sin su tenacidad, dedicación y amor nada hubiera sido posible.

A mi adorado padre, Enrique. Por confiar en mi capacidad y ayudarme a cumplir mis sueños. Por sus consejos y escucharme siempre con paciencia. Por las sonrisas de satisfacción por cada objetivo alcanzado.

A mi querida mamita Alicia, por estar aún no estando. Por acompañarme en cada paso y dejar una huella imborrable en mi vida y en mi corazón.

Este logro no es mío sino de ustedes

Con amor,

Mía

A la mujer más importante en mi vida, mi madre, Ana. Eternamente agradecido con todo el sacrificio y amor que me brindo desde el primer momento, este logro es uno más que la vida me permite compartir contigo.

Y a todas aquellas personas que estuvo para mí a lo largo del duro camino, pues son los tesoros más preciados que tengo y esto también se lo dedico a ustedes. Finalmente, a Shaku, el pug, el mejor compañero que me dio la vida.

Gracias totales,

Diego

A nuestro querido profesor José Taquía, por sus enseñanzas, orientación, tiempo y dedicación en la realización de este trabajo desde el primer día.

## REFERENCIAS

Alaiz-Rodriguez, R. y Parnell, A. (2020). Un enfoque de aprendizaje automático para la evaluación de la calidad de la carne de cordero utilizando espectros FTIR. Acceso IEEE.

Athey, S., Tibshirani, J. y Wager, S. (2019). Bosques aleatorios generalizados. Anales de Estadística.

Bovo, M., Agrusti, M., Benni, S., Torreggiani, D. y Tassinari, P. (2021). Modelado forestal aleatorio de la producción de leche de vacas lecheras en condiciones de estrés por calor. Animales.

Cunningham, P. y Delany, S. (2021). k-Clasificadores de vecinos más cercanos: un tutorial. ACM Computing Surveys, 25.

Denholm, S. J., Brand, W., Mitchell, A. P., Wells, A. T., Krzyzelewski, T., Smith, S. L., . . . Coffey, M. P. (2020). Predicción del estado de la tuberculosis bovina de las vacas lecheras a partir de datos espectrales del infrarrojo medio de la leche utilizando aprendizaje profundo. Revista de Ciencia Láctea.

Dórea, J., Rosa, G., Weld, K. y Armentano, L. (2018). Extracción de datos de la espectroscopia infrarroja de leche para mejorar las predicciones de ingesta de alimento en vacas lecheras lactantes. Revista de Ciencia Láctea.

- Feng, L., Zhu, S., Chen, S., Bao, Y. y él, Y. (2019). Combinación de espectroscopia de infrarrojo medio por transformada de Fourier con métodos quimiométricos para detectar adulteraciones en leche en polvo. *Sensores*.
- Frizzarin, M., O'Callaghan, T., Murphy, T., Hennessy, D. y Casa, A. (2021). Aplicación de métodos de aprendizaje automático a la leche de espectros de infrarrojo medio para la discriminación de leche de vaca de pastos o dietas de raciones mixtas totales. *Revista de Ciencia Láctea*.
- Fuentes, G., Ruiz, R., Sánchez, J., Ávila, D. y Escutia, J. (2013). Análisis microbiológico de leche de origen orgánico: atributos deseables para su transformación. *Agricultura Sociedad y Desarrollo*, 14.
- Fuentes, S., Viejo, C., Cullen, B., Tongson, E., Chauhan, S. y Dunshea, F. (2020). Inteligencia artificial aplicada a una granja lechera robótica para modelar la productividad y calidad de la leche en base a datos de vacas y parámetros ambientales diarios. *Sensores*.
- Gonzales, J., Portocarrero, S. y Abanto, M. (2019). Calidad de los derivados lácteos producidos en la Región Amazonas, Perú. *Revista Científica UNTRM: Ciencias Naturales e Ingeniería*, 6.
- Huang, Y., Doh, I. y Bae, E. (2021). Diseño y validación de una nariz electrónica portátil basada en machine learning. *Sensores*.
- Kang, S. (2021). k-Aprendizaje del vecino más cercano con redes neuronales gráficas. *Matemáticas*, 12.
- Keerthana, S. y Santhi, B. (2020). Encuesta sobre aplicaciones de nariz electrónica. *Revista de Ciencias de la Computación*.
- Kumar, I., Rawat, J. y Mohd, N. (2021). Oportunidades de Inteligencia Artificial y Machine Learning en la Industria Alimentaria. *Revista de Calidad de los Alimentos*.
- Mengucci, C., Rabiti, D., Urbinati, E., Picone, G., Romano, R., Aiello, A., . . . Capozzi, F. (2021). Detección de cuajada congelada en queso mozzarella de búfala DOP a través de información sobre su estructura supramolecular adquirida por experimentos de relajación 1H TD-NRM. *Ciencias aplicadas*.
- Mera, L. y Ochoa, J. (2021). Redes neuronales convolucionales para la clasificación de componentes independientes de rs-fMRI. *Atención Primaria* , 19.
- Mu, F., Gu, Y., Zhang, J. y Zhang, L. (2020). Identificación de fuentes de leche y estimación de la calidad de la leche utilizando técnicas electrónicas de nariz y aprendizaje automático. *Sensores*.
- Neto, H., Tavares, W., Ribeiro, D., Alves, R., Fonseca, L. y Campos, S. (2019). Sobre la utilización del aprendizaje profundo y conjunto para detectar la adulteración de la leche. *Minería de biodatos*.
- Pathot, Y. D. (2019). Prácticas Higiénicas y Calidad Bacteriológica de Leche: Una revisión . *Revista Internacional de Investigación-Granthaalayah* , 16.
- Satoła, A. y Bauer, E. (2021). Predicción de cetosis subclínica en vacas lecheras utilizando técnicas de aprendizaje automático. *Animales*.
- Sarmiento-Ramos, J. L. (2019). Aplicaciones de redes neuronales y profundas. *UIS Ingenierías*, 18.

- Schonlau, M. y Yuyan Zou, R. (2020). El algoritmo de bosque aleatorio para estadística. Diario Stata.
- Sharif, A., Abbasi, Q., Arshad, K., Ansari, S., Ali, M., Kaur, J., . . . Imran, M. (2021). El aprendizaje automático permitió la detección de contaminación de alimentos utilizando rfid y el sistema de Internet de las cosas. Journal of Sensor and Actuator Networks.
- Sharma, A., Jain, A., Gupta, P. y Chowdary, V. (2021). Aplicaciones de aprendizaje automático para la agricultura de precisión: una revisión exhaustiva. Acceso IEEE.
- Shen, W., Cheng, F., Zhang, Y., Wei, X., Fu, Q. y Zhang, Y. (2020). Reconocimiento automático de comportamientos relacionados con la ingesta de vacas lecheras basado en la aceleración triaxial. Procesamiento de la información en la agricultura.
- Stefas, D., Gyftokostas, N., Kourelis, P., Nanou, E., Kokkinos, V., Bouras, C. y Couris, S. (2021). Un método basado en láser para la detección de la adulteración de la miel. Ciencias aplicadas.
- Takruri, M., Abubakar, A., Alnaqbi, N., Shehhi, H., Jallad, A. y Bermak, A. (2020). DoFP-ML: Un enfoque de aprendizaje automático para el monitoreo de la calidad de los alimentos utilizando un sensor de imagen de polarización DoFP. Acceso IEEE.
- Yan, J., Jin, M., Xu, Z., Chen, L., Zhu, Z. y Zhang, H. (2021). Reconocimiento de líquido de suspensión basado en patrones de moteado mediante aprendizaje profundo. IEEE Photonics Journal.

## ANEXO.

### Datos del artículo publicado

- **Nombre del artículo:** Machine Learning Applied to Milk Sample Classification
- **Autores:** Mía León Loyola y Diego Ossa De la Cruz.
- **Co autor(es):** José Antonio Taquíá Gutiérrez

### Presentación en congreso

- **Nombre del congreso:** First Australian Conference on Industrial Engineering and Operations Management
- **Organizador:** IEOM Society
- **Sede:** Sydney, Australia
- **Año:** 2022
- **Pp:** 1-9
- **Enlace web donde se encuentra publicado el artículo (identificador DOI, ISBN, ISSN o equivalentes):** <https://doi.org/10.46254/AU01.20220511>

# OssaLeon

---

## INFORME DE ORIGINALIDAD

---

6%

INDICE DE SIMILITUD

4%

FUENTES DE INTERNET

1%

PUBLICACIONES

2%

TRABAJOS DEL ESTUDIANTE

---

ENCONTRAR COINCIDENCIAS CON TODAS LAS FUENTES (SOLO SE IMPRIMIRÁ LA FUENTE SELECCIONADA)

---

1%

★ [repositorio.ulima.edu.pe](http://repositorio.ulima.edu.pe)

Fuente de Internet

---

Excluir citas

Activo

Excluir coincidencias

Apagado

Excluir bibliografía

Activo