

# Implementación de Procesos KDD (Knowledge Discovery in Databases) para la caracterización de la calidad de leche bovina en función de la composición e higiene sanitaria

---

---

Año  
2017

Autor  
Cabrera, Gabriela Pilar

Este documento está disponible para su consulta y descarga en el portal on line de la Biblioteca Central "Vicerrector Ricardo Alberto Podestá", en el Repositorio Institucional de la **Universidad Nacional de Villa María**.

#### CITA SUGERIDA

Cabrera, G.P.; [et al.] (2017). *Implementación de Procesos KDD (Knowledge Discovery in Databases) para la caracterización de la calidad de leche bovina en función de la composición e higiene sanitaria*. Villa María: Universidad Nacional de Villa María



## Implementación de Procesos KDD (Knowledge Discovery in Databases) para la caracterización de la calidad de leche bovina en función de la composición e higiene sanitaria.

**Director:** CABRERA Gabriela Pilar

**Integrantes:** MARCO Laura Angélica, ASINARI Marianela Cristina, CORNEJO Anabella Patricia, PAVÓN Sandra Marisel, DEFAGOT María Verónica Raquel.

### Informe Académico

En este proyecto de investigación se propone el análisis, aplicación y pertinencia del proceso de KDD (Knowledge Discovery in Databases) para la caracterización de la calidad de leche bovina cruda -en aspectos Higiénicos-Sanitarios, Composicional y Adulteración- y sus variaciones estacionales; registradas durante el año 2014 de cinco empresas lácteas del departamento Río II, Provincia de Córdoba. Para ello, durante el año 2016 se cumplimentaron las etapas de limpieza e integración, obtención del Data warehouse y en este momento se trabaja en la evaluación y presentación de patrones.

Específicamente, se completó el pre-procesamiento de los datos provistos por los usuarios (Laboratorio de Diagnóstico Villa María), lo que posibilitó con la asistencia de expertos en Calidad de Leche Bovina, comprender el dominio del caso de estudio.

Un punto muy importante relacionado con esta información tiene que ver con la modalidad de registración de los datos, son cuatro registraciones mensuales por empresa láctea (se tienen cinco industrias codificadas con A, B, C, D y E) durante los 12 meses del año 2014.

Se dio cumplimiento a la obtención y depuración del Data Warehouse. La referida base de datos (Figura 3), consta de 1976 registros y con 14 variables de análisis (entre las cuales se debieron realizar categorizaciones específicas); conformándose por tanto una tabla de 1976 filas x 14 columnas.

En el momento actual, el equipo de investigación se aboca a la evaluación de la calidad de los datos generados y la identificación de patrones en los mismos, a través de un análisis exploratorio; mediante la aplicación de software estadísticos IBM Spss, InfoStat y el lenguaje y entorno de programación R.

A partir de este análisis exploratorio, se encontró un porcentaje muy bajo de muestras "sanas", en tanto que más del 50% de las muestras se manifiestan con problemas o su condición es grave. Esta situación es un indicador de atención respecto del objetivo de

determinar la calidad de la leche pequeñas y medianas industrias (tambos). Por otra parte, se observa que una de las industrias parece estar en mejores condiciones que las restantes respecto del nivel de células somáticas. Sigue entonces, analizar las cuestiones que favorecen dicha situación.

### Producción científica relevante

Por otra parte, cabe destacar que el presente proyecto surge como respuesta a la demanda del LABORATORIO DE DIAGNOSTICO VETERINARIO VILLA MARIA, en relación a la necesidad de caracterizar la composición y calidad microbiológica de leche cruda de tambos bovinos. Cabe señalar, que estos tambos resultan pequeñas o medianas empresas y en este sentido interesa identificar si la calidad de leche se ve afectada por el tipo de emprendimiento.