

METODOS DE CLASIFICACIÓN Y PREDICCIÓN EN CONTEXTOS DE BIG-DATA A DOS Y TRES MODOS

CÓDIGOS

UNR 80020180300041UR
Programa de Incentivos

FECHAS

Inicio 2019
Finalización 2022

RADICACIÓN

Instituto Instituto de Investigaciones Teóricas y Aplicadas
Escuela Escuela de Estadística

CAMPO

Disciplina Científica Estadística – Métodos Estadísticos Específicos
Especialidad Análisis Multivariados – Data Mining – Big Data
Campo de Aplicación Producción Vegetal – Desarrollo socioeconómico y servicios

RESUMEN

Los grandes avances producidos en las tecnologías de información, han generado nuevos desafíos en el área del análisis de datos, requiriendo de metodología que permita descubrir y entender más allá de lo que las herramientas tradicionales plantean. El análisis estadístico multivariado aplicado, conocido como Data Mining o Big Analytics, permite, según una de sus definiciones, la extracción de información valiosa y conocimiento procesable que está implícito en grandes cantidades de datos. El análisis de los mega-datos suele tener variados propósitos: predicción, produciendo una apropiada categorización de nuevos objetos en niveles en función de sus atributos; clasificación o segmentación, agrupando objetos de modo que dentro de cada grupo exista cierta homogeneidad o parecidos y a su vez diferencias marcadas entre los grupos; descubrir outliers, es decir objetos inusuales dentro de un conjunto; encontrar modelos locales, encontrando pequeños subgrupos de objetos que tengan fuertes relaciones entre ellos mismos; o identificar variables latentes o no observables que son las que originan las relaciones entre grupos de variables observadas, enfocando en ocasiones las relaciones de causalidad. Las técnicas que se aplican se dividen en dos áreas principales: análisis supervisado y análisis no supervisado. El análisis multivariado supervisado tiene como objetivo, predecir una variable o campo de interés -llamado respuesta- en función de un gran conjunto de variables predictoras, y a su vez, entender cómo es la relación entre las mismas y la respuesta. Cuando la variable respuesta que se desea predecir es cuantitativa continua, se suele decir que se está ante un problema de Regresión, mientras que si la respuesta es cualitativa, es decir, se pretende determinar una categoría de pertenencia, se habla de un problema de Clasificación. Este proyecto plantea avanzar en la investigación de métodos supervisados y no supervisados, explorando nuevas propuestas de métodos multivariados de clasificación y de predicción apropiados para el análisis de big-data, a dos y tres vías, verificando sus propiedades y posibilidades de aplicación, tanto en casos reales, como a través de intensivos estudios por simulación.

PALABRAS CLAVE

Data Mining Regresiones LASSO Métodos de clasificación

GRUPO DE INVESTIGACIÓN

Apellido y Nombre	Dedicación	Función	Categoría Incentivos	Participación
Quaglino, Marta Beatriz	Exclusiva	Directora	I	2019-2022
Vitelleschi, María Susana	Simple	Codirectora	III	2019-2022
Allasia, María Belén	Simple	Integrante	V	2019-2022
Angelini, Julia	Simple	Integrante	-	2019-2022
Del Médico, Ana Paula	Simple	Integrante	-	2019-2022
Faviere, Gabriela Soledad		Integrante		2019-2022
Kovalesvsky, Leandro Oscar	Semiexclusiva	Integrante	IV	2019-2022
Macat, Paula	Simple	Integrante	V	2019-2022
Severino, Liliana Carmen	Semiexclusiva	Integrante	IV	2019-2022

RESOLUCIONES DEL CONSEJO SUPERIOR

Acreditación		Continuidad 2020		Continuidad		Continuidad		Aprobación Informe Final	
C.S.Nº	En trámite	C.S.Nº		C.S.Nº		C.S.Nº		C.S.Nº	
Fecha		Fecha		Fecha		Fecha		Fecha	

FINANCIAMIENTO

Año	2019	Año	2020	Año	2021	Año	2022
Monto		Monto		Monto		Monto	
C.S.Nº		C.S.Nº		C.S.Nº		C.S.Nº	
Fecha		Fecha		Fecha		Fecha	