

MÉTODOS MULTIVARIADOS DE PROYECCIÓN EN DATA MINING.

CÓDIGOS		FECHAS	
UNR	1ECO127	Inicio	2011
Programa de Incentivos	19/E320	Finalización	2014

RADICACIÓN	
Instituto	Instituto de Investigaciones Teóricas y Aplicadas
Escuela	Estadística

CAMPO	
Disciplina Científica	Matemática
Especialidad	Estadística
Campo de Aplicación	

RESUMEN

El análisis de datos en la actualidad, presenta características que eran impensables hace una par de décadas atrás.

La posibilidad de almacenaje de datos en forma automática por medios informáticos y la capacidad de las computadoras de realizar operaciones complejas con extrema rapidez, han puesto a los analistas frente a nuevos desafíos. Estas particularidades que aparecen como potenciales ventajas al momento de recoger información, pueden acarrear inconvenientes al momento de extraer la información que contienen, apareciendo particularidades como datos con errores difíciles de detectar en el gran conjunto, datos faltantes, variables colineales, variables sin información relevante, mediciones multivariadas de distinta longitud, registro de "variables" no convencionales como textos o imágenes, matrices de covariancia con rango menor a su dimensión, etc. El amplio conjunto de técnicas estadísticas desarrolladas y utilizadas para interpretar los fenómenos que tales datos representan, son técnicas multivariadas que se han reunido bajo el nombre de Data Mining, haciendo alusión a las actividades mineras, capaces de extraer materiales con valor económico de yacimientos donde están escondidos.

El presente proyecto plantea el avance en el estudio, generación de conocimiento y aplicación de estas técnicas, partiendo de la búsqueda de soluciones a problemáticas específicas, en las que miembros del equipo de investigación están trabajando.

En principio se proponen tres áreas de aplicación, que requieren de métodos particulares: la de evaluación de calidad de vacunas destinadas a bovinos, que evitarían una infestación que constituye una causa importante de pérdidas económicas en la actividad ganadera en nuestro país (utilizando Structural Equation Models-SEM), la del control de calidad de procesos industriales (a través de Multivariate Statistical Process Control-MSPC), y la de evaluación de publicidades en investigación de mercados (por medio de técnicas de Multidimensional Scaling-MDS).

PALABRAS CLAVE			
Data Mining	SEM	MSD	MSPC

GRUPO DE INVESTIGACIÓN

Apellido y Nombre	Dedicación	Función	Categoría Incentivos	Participación
Quaglino, Marta Beatriz	Exclusiva	Director	II	2011-2014
Vitelleschi, María Susana	Exclusiva	Codirector	IV	2012-2014
Vitelleschi, María Susana	Exclusiva	Integrante	IV	2011
Ballarini, Nicolás	Simple	Integrante	-	2011-2014
Kovalesky, Leandro	Exclusiva	Integrante	-	2011-2014
Macat, Paula	-	Integrante	-	2014
Zárate, Melina	Simple	Auxiliar	-	2011-2013

RESOLUCIONES DEL CONSEJO SUPERIOR

Acreditación		Continuidad 2012		Continuidad 2013		Continuidad 2014		Aprobación Informe Final	
CS. N°	140/2011	CS. N°	396/2012	CS. N°	-	CS. N°	494/2014	CS. N°	091/2017
Fecha	08/09/2011	Fecha	05/06/2012	Fecha	-	Fecha	31/07/14	Fecha	28/03/17

FINANCIAMIENTO

Año	2011	Año	2012	Año	2013	Año	2014
Monto	\$5.177,72	Monto	7.598,87	Monto	\$4.444	Monto	\$4.299
C.S.N°	949/2013	C.S.N°	430/2015	C.S.N°	26.499	C.S.N°	26.499
Fecha	05/12/2013	Fecha	18/06/2015	Fecha	01/08/2017	Fecha	01/08/2017