

**RINCE**

**Revista de Investigaciones del Departamento de Ciencias Económicas  
de La Universidad Nacional de La Matanza**

***Comunicación Científica***

**Dossier**

***I Jornada de Investigación Interdepartamental:***

***"25 Años de Desarrollo e innovación en el conocimiento"***

**Universidad Nacional de La Matanza**

- 1. Título de la Ponencia:** Sistemas Grises y Conjuntos Rugosos en el tratamiento de información con incertidumbre
- 2. Área y tema al cual pertenece:** Línea temática N° 9: Desarrollos en Administración, Economía y Contabilidad
- 3. Nombre de la Jornada, Seminario, Congreso u otro tipo de evento científico:** 1ra Jornada de Investigación Interdepartamental: "25 Años de Desarrollo e innovación en el conocimiento"
- 4. Lugar y fecha de realización:** Universidad Nacional de La Matanza, San Justo. Provincia de Buenos Aires. Argentina, 15 de septiembre de 2014.
- 5. Nombre/s y Apellido/s del/los autor/es:** Osvaldo Galardo, Jorge Barreto, Sebastian Malagrino, Marcelo Perisse, Juan Roger, Patricia Fusco
- 6. Domicilio particular y/o laboral:** Departamento de Ciencias Económicas. Florencio Varela 1903 San Justo, Buenos Aires.
- 7. Teléfono particular, laboral y/o celular:** 4480-8954.
- 8. Dirección de correo electrónico particular y/o comercial:**  
[ogalardo@gmail.com](mailto:ogalardo@gmail.com)
- 9. Nombre de la Institución que aceptó el trabajo:** Comité Evaluador. Secretaría de Ciencia y Tecnología. Universidad Nacional de La Matanza.
- 10. Link de acceso directo al evento y/o institución responsable de la publicación:**  
<http://cyt.unlam.edu.ar/index.php?seccion=17&idArticulo=649>

### **11. Resumen:**

El rápido desarrollo y la extensa aplicación de la tecnología de bases de datos, unido a la gran cantidad de datos disponibles para el análisis en distintos ámbitos de las ciencias puras y aplicadas, ha convertido la minería de datos (data mining) y el descubrimiento de conocimiento (knowledge discovery) en áreas de intensa investigación. El conocimiento se ha convertido en un componente imprescindible en la industria, la banca e incluso en la vida social. En particular, los sectores económicos -tanto en las áreas de las manufacturas como de las finanzas- requieren tomar decisiones sobre gran cantidad de información para obtener ventajas competitivas. Sin embargo, la información es frecuentemente imprecisa debido a que contiene gran cantidad de datos pero es cualitativamente pobre. La extracción de información comprensible a partir de grandes volúmenes de datos hace uso del proceso KDD (Knowledge Discovery in Databases) que opera para obtener información subóptima pero económica y confiable para la toma de decisiones, en los casos en que la información es incompleta, inexacta o inconsistente. Esta investigación se centra en dos técnicas específicas: la teoría de conjuntos aproximados o rugosos (Rough Set Theory o RST) y la teoría de sistemas grises (Grey Systems Theory o GST), y en la hibridación de ambas: Grey Rough Set Theory o GRST. Se exhiben las propiedades fundamentales de las tres teorías y se evalúa la conveniencia de generalizaciones y ampliaciones de dichas teorías para el tratamiento de información incierta.

**12. Palabras claves:** Información incierta; conjuntos aproximados o rugosos y sistemas grises; hibridación gris- rugosa.

## **Desarrollo de la ponencia**

### **13. Planteo del problema:**

Los procesos micro y macroeconómicos, las actividades de marketing y –en general- las de gerenciamiento, se corresponden con problemas complejos de decisión para los cuales la información disponible es frecuentemente incompleta. El tratamiento de este tipo de problemas en condiciones de incertidumbre requiere el aporte de distintas disciplinas y en particular de las tecnologías “soft computing” que se apoyan en distintos tratamientos matemáticos, de los cuales son bien conocidos las redes neuronales, los algoritmos genéticos y la lógica difusa. Pero en las últimas décadas se ha desarrollado dos líneas promisoras: la Teoría de conjuntos rugosos – introducida por Z. Pawlak en 1982- y –también en 1982- la Teoría de Sistemas Grises por D. Julong; con el aporte de ambas líneas se han impulsado los sistemas híbridos para el tratamiento de la incertidumbre en la información que conectan las distintas tecnologías citadas. Esta integración alcanza la teoría de Probabilidades y los desarrollos en Meteorología.

### **14. Breve descripción del proyecto de investigación:**

El objetivo de esta investigación es una puesta al día de los sistemas grises, los conjuntos rugosos y –complementariamente- algunas formas híbridas que se aplican en soft computing y son relevantes para el tratamiento de la información incompleta, en especial aplicada a la Economía, Marketing y Gerenciamiento en general. Se realizó una extensa revisión y selección bibliográfica de investigaciones específicas incluyendo el análisis comparativo de ambos recursos y su integración en sistemas híbridos.

Se espera que esta investigación contribuya en el establecimiento de bases para la formación de recursos humanos y el desarrollo de programas de computación, basados en sistemas híbridos, que agilicen y mejoren la eficiencia de la toma de decisiones en distintos ámbitos económicos.

## **15. Conceptos teóricos que sirven de marco a la investigación:**

### a) Rough Set Theory:

RST (Rough Set Theory) fue propuesta por Zdzislaw Pawlak en 1982 como una nueva y potente herramienta matemática para tratar información vaga o imprecisa. La idea básica es la clasificación de objetos (concretos o abstractos, como señales, eventos, procesos, entidades) incluyendo en el mismo conjunto a objetos que posean diferencias menores entre sí. Por lo tanto, objetos que portan la misma o similar información son indiscernibles (respecto de la información disponible). Las unidades de distintos conjuntos indiscernibles son los conjuntos de elementos que componen el universo de la información disponible.

RST establece definiciones precisas (es decir, lógico-matemáticas) de los componentes conceptuales básicos y obtiene, en su modelización, los teoremas que facilitan las aplicaciones y desarrollos específicos. En RST se define la Relación de Indiscernibilidad en el universo de datos; dicho universo está formado por subconjunto de objetos y propiedades. Con relación a una propiedad y dado un objeto perteneciente a un subconjunto del universo puede darse: A) que el objeto pertenezca absolutamente (o, con la misma imprecisión semántica: 'seguramente') al subconjunto; B) que el objeto no pertenezca absolutamente al subconjunto; C) que el objeto pueda pertenecer, o no, al subconjunto. Debido al punto C) queda claro que el subconjunto en cuestión es impreciso y para manejarlo se requieren dos conceptos centrales de RST: aproximación superior y aproximación inferior, ambas definidas en conexión con la relación de indiscernibilidad. La aproximación inferior es la unión de todos los objetos que pertenecen absolutamente al subconjunto y se conoce como región positiva del subconjunto. La aproximación superior es el menor conjunto de elementos que pueden pertenecer aparentemente al subconjunto en análisis. La región negativa contiene los objetos que no pertenecen al subconjunto. La diferencia entre la aproximación superior y la aproximación inferior es la frontera o límite que no puede ser evaluada respecto del atributo en cuestión, porque allí los objetos no pueden clasificarse como pertenecientes, o no, al subconjunto. Si la frontera es vacía, entonces el subconjunto es exacto en términos de indiscernibilidad; si la

frontera no es vacía, el subconjunto es rugoso (áspero, impreciso) considerando dicha relación. El sistema de información (o de representación) sobre el que opera RST está formado por el universo de subconjuntos de objetos, el conjunto de propiedades o atributos y la función cuyo dominio es el producto cartesiano de subconjuntos del universo por subconjuntos de propiedades y cuya imagen es el conjunto formado por los subconjuntos en los que se asignan objetos a la propiedad dada. Existen formas alternativas basadas en relaciones distintas a la de indiscernibilidad (que es una relación de equivalencia): A) en base a la relación de similaridad, que no genera particiones del universo de objetos y retiene la reflexividad pero no la simetría ni la transitividad; B) en base a relaciones binarias reflexivas y transitivas

#### B) Grey System Theory:

El mismo año (1982) que Pawlak propone RST, el investigador chino Julong Deng presenta su teoría de Sistemas Grises (GST) que, en poco menos de una década, tiene una aplicación exitosa en gran variedad de campos como agricultura, hidrología, historia, ecología, medicina, geografía, ciencias de los materiales, protección biológica, sistema judicial, industria, temas militares y en distintas áreas de la Economía. En China se aplicó exitosamente en economía agrícola, estimación de efectos económicos, modelos diagnósticos en medicina, protección biológica y meteorología.

Según el autor de GST, 'grey' es un término que –relativo a la información que provee un sistema- significa 'pobre', 'incompleta', 'incierto'. En este sentido, gran cantidad de sistemas complejos son grises porque varios de sus parámetros son inciertos o desconocidos; un ejemplo concreto es el cuerpo humano porque varios parámetros de su medio interno no se conocen o la información disponible es pobre.

Cuando la información es pobre, imprecisa o incompleta, GST provee alternativas que para procesos y modelos estocásticos, análisis de regresión o análisis de series de tiempos. Dado que los métodos estadísticos tradicionales requieren funciones de distribución de probabilidad y en el mundo real tales funciones no

están usualmente disponibles, la modelización en variables grises puede satisfacer las necesidades de conocimiento.

GST comprende varias áreas: matemática gris, espacios relacionales grises, control gris, generación de espacios grises y modelización gris, entre otros. La modelización gris en una variable (GM 1,1) y el Análisis Relacional Gris (GRA) son dos de los métodos usados frecuentemente.

La unidad o célula de un sistema gris es el número gris, que expresa el carácter del comportamiento del sistema gris. De un número gris se conoce el rango de valor aunque se desconoce el valor preciso. Los sistemas con información completa son sistemas “blanqueados” (whitening) o “blancos”, los que carecen completamente de información son sistemas “negros” y los que contienen información pobre o incompleta son sistemas grises.

Dado que el número gris expresa la calidad de la información analizada, una cuestión fundamental de GST consiste en aumentar la precisión del valor original; es decir, en hacer que el número gris del sistema se logre “blanquear”.

### C) Hibridación RST-GST:

RST y GST son contemporáneas en su origen y ambas efectivas en computación soft para tratar con sistemas de información que contienen datos imprecisos. Por supuesto, no son las únicas y un ejemplo muy próximo es el sistema fuzzy o borroso. En casi todos los casos los términos utilizados para denominar estas distintas teorías son sólo nombres que permiten identificarlas y en su elección se privilegia la idea de imprecisión asociada a alguna intuición generalmente sensorial: rugoso/áspero (táctil), borroso (visual) o gris (color). Esta costumbre ya es habitual en otros ámbitos científicos como en Física de Partículas donde se recurre a términos como ‘quark up’ y ‘quark down’ que son simples etiquetas para diferenciar distintos objetos físicos. Aunque dicha elección es creativa, en el caso específico de RST y GST introduce cierta confusión inicial porque todas se refieren a información imprecisa pero el término empleado en cada caso (rough, grey) no da ninguna pista sobre el enfoque de cada una. Desde ya, esto se soluciona analizando las estructuras axiomáticas respectivas lo cual es imprescindible no sólo para lo anterior sino, y principalmente, para evaluar la

posibilidad y la conveniencia del uso asociado (hibridación) a fin de explotar lo mejor de cada teoría. Este último punto es particularmente delicado porque involucra varios aspectos a tener en cuenta, entre ellos: cómo trata la imprecisión del sistema informativo cada teoría; en ese caso, si es viable una aplicación sucesiva de técnicas o es aconsejable, en cambio, una integración; cuántas técnicas conviene hibridar a la vez, por citar algunos aspectos tratados frecuentemente. Sin embargo, las técnicas de hibridación son una opción a tener en cuenta porque cada teoría tiene aplicaciones específicas y ninguna –hasta ahora- resuelve todos los problemas. En la actualidad están disponibles numerosas aplicaciones donde se emplean variadas hibridaciones de teorías incluyendo distintas versiones de las mismas hibridaciones para usos específicos, que pueden resultar equivalentes en resultados. Éste es el caso de la hibridación que se analiza en este trabajo (RST-GST), para la cual se han descrito distintas asociaciones como: aproximaciones rugosas basadas en conjuntos grises y funciones de blanqueo; o RST basada en álgebra de intervalos que soportan operaciones con retículos grises. También se han descrito equivalencias entre hibridaciones RST-GST con extensiones de RST y la asociación RST-GST con el análisis de conceptos formales (FCA)

### **16.Aspectos relevantes de la investigación:**

En el caso particular de GRST, la hibridación ha resultado satisfactoria en el tratamiento de información compleja, variable e incierta particularmente en Economía y Finanzas. En este aspecto, es importante destacar que GRST incorpora la experiencia de los analistas de datos y algunas extensiones de la teoría, como en los modelos basados en relaciones de tolerancia (matemáticamente: reflexivas y simétricas) se emplean valores de umbral de información, superiores e inferiores, que mejoran los resultados.

**17. Aportes alcanzados en relación con el problema planteado:**

Aunque se han realizado análisis comparativos entre distintas técnicas, la tendencia actual parece orientarse a la diversificación y enriquecimiento teórico en las técnicas individuales y en las hibridaciones.