

NUEVAS TECNOLOGÍAS PARA EL MARKETING PÚBLICO: MINERÍA DE DATOS Y TÉCNICAS BIO-INSPIRADAS

Enrique López González (ddeelg@unileon.es)
Cristina Mendaña Cuervo (ddecmc@unileon.es)

Departamento de Dirección y Economía de la Empresa
Universidad de León
Facultad de Ciencias Económicas y Empresariales
Campus de Vegazana, s/n
E-24071 León

RESUMEN

El propósito del presente trabajo es presentar la utilidad de aplicar tecnologías emergentes en el marketing público, especialmente relativas a la Minería de Datos para llevar a cabo criterios de segmentación. A este respecto, dado que el conocimiento que se tiene del comportamiento de las variables de interés en el proceso de toma de decisiones es impreciso, se debe incluir entonces la noción de nivel de presunción, lo que plantea la necesidad de un acercamiento a aquellas herramientas matemáticas que permitan procesar esa información y trabajar con valoraciones subjetivas. De esta forma, se avizora el interés por las denominadas “Heurísticas Bio-inspiradas” (Redes Neuronales, Algoritmos Genéticos y Sistemas de Hormigas, entre otras).

PALABRAS CLAVE: *Segmentación, Minería de Datos, Heurísticas bio-inspiradas*

1. INTRODUCCIÓN

A los efectos del presente trabajo se parte de una concepción del marketing público como conjunto de actividades cuyo objetivo es el diseño, implementación y control de programas destinados a satisfacer las necesidades de los usuarios, personas naturales o jurídicas, de los servicios brindados por el sector público, mediante el diseño adecuado del servicio, de la distribución, de la promoción, del personal, de la presencia física, de los procesos y, eventualmente, de los precios.

Por tanto, se trata de una definición de marketing como función administrativa o gerencial, es necesario analizar algunos aspectos de esta definición propuesta. Primero, es una actividad gerencial, esto es, tal actividad es responsabilidad de quienes actúan como gerentes de las instituciones públicas, ya sea Ministros, Subsecretarios, Directores, Jefes de Sección o cualquier otra denominación. En segundo lugar, el objetivo es la aceptación del servicio público ofrecido a un grupo de personas, naturales o jurídicas. La orientación a las necesidades de los consumidores está presente en esta definición. En tercer lugar, las variables controlables por el gerente público no se tienen por que restringirse a las clásicas del marketing privado. Así, por ejemplo a las cuatro variables tradicionales (producto/servicio, precio, canales de distribución y promoción) cabe agregar otras propias del marketing de servicios: las personas que brindan el servicio, la presencia física (instalaciones) en las cuales se brinda el servicio y los procesos utilizados para brindar este servicio. Todos elementos controlables que permiten satisfacer las necesidades de los consumidores o usuarios del servicio.

En este sentido, una de las herramientas más poderosas del marketing es la idea de segmentar el mercado y seleccionar segmentos antes de elaborar el plan de marketing. En efecto, en el marketing privado existe gran variedad de criterios de segmentación: geográficos, demográficos, socio económicos, psicográficos y conductuales, mientras que en el marketing público es igualmente necesario segmentar los mercados, pero los criterios utilizados son bastante más limitados, si bien los criterios geográficos y demográficos son ampliamente utilizados. Así, por ejemplo, cabe señalar la Psicografía, esto es, el empleo sistemático de los constructos relativos a la actividad, intereses y opiniones para analizar y explicar cuantitativamente el comportamiento de comunicación, compra y consumo por personas de las marcas, productos y grupos de productos. Es un método para definir el estilo de vida de una persona en términos mensurables. Hoy en día es considerada una de las técnicas más novedosas, interesantes y prometedora para escoger un mercado meta. También, el criterio de conducta es igualmente utilizado en el marketing privado que en el marketing público, especialmente aquel que dice relación con la cantidad o frecuencia de uso del servicio: así por ejemplo el Registro Civil brinda servicios especiales a aquellas instituciones en que un gran número de trabajadores debió renovar su cédula de identidad.

Por consiguiente, cabe afirmar que los criterios de segmentación utilizados en marketing privado y en el marketing público son muy similares. Las principales diferencias se encuentran en la forma como se utilizan los criterios socio-económicos, y la absoluta falta de utilización de criterios psicográficos en las entidades públicas.

De acuerdo con lo anterior, el objetivo principal del presente trabajo es presentar la utilidad de aplicar nuevas tecnologías en el marketing público, para lo cual en el siguiente apartado se analizará el papel que puede desempeñar a este respecto la Minería de Datos y como esta a su vez puede desarrollarse aplicando tecnologías basadas en la naturaleza, a cuya consideración se dedicará la segunda parte del trabajo.

La novedad aportada en el presente trabajo radica en que la emergencia de la aplicación de dichas heurísticas bio-inspiradas se basa en que si el conocimiento que se tiene del comportamiento de las variables de interés en el proceso de toma de decisiones es impreciso, se debe incluir entonces la noción de nivel de presunción, lo que plantea la necesidad de un acercamiento a aquellas herramientas matemáticas que permitan procesar esa información y trabajar con valoraciones subjetivas. Los números borrosos han sido creados para reflejar la vaguedad de la percepción humana y con ella la noción de presunción. De esta forma, se avizora el interés por las denominadas “Técnicas Inteligentes” (Lógica Borrosa, Redes Neuronales y Algoritmos Genéticos, entre otras) cuyas características distintivas de las técnicas operativas convencionales radican en dos aspectos básicos, a saber: (i) en que su principal objetivo es sacar provecho de la tolerancia que conlleva la vaguedad y la incertidumbre propias de los problemas mal estructurados, para alcanzar resultados que sean robustos y comprensibles para los decisores, y (ii) que los algoritmos operativos de las mismas simulan los mecanismos de la

naturaleza para resolver problemas, siendo ésta una fuente inagotable de ideas para el desarrollo técnico y científico.

2. LA MINERÍA DE DATOS

La Minería de Datos pretende obtener visiones en profundidad de los datos corporativos que no son fácilmente detectables. De hecho, más que analizar los resultados de la actividad, permiten modelizarla construyendo patrones o categorías que la identifiquen, respondiendo a las necesidades de información del tipo ¿qué hay en los datos de interés? o ¿qué podría ocurrir en un futuro? en base al descubrimiento de tendencias o agrupaciones interesantes de datos. De hecho, las herramientas enmarcadas bajo la denominación de Minería de Datos (MD), permiten no sólo el análisis de información que tradicionalmente ha venido siendo realizado por los Sistemas de Soporte a la Decisión (DSS), sino, y esto es lo realmente importante y diferencial, el planteamiento y descubrimiento automático de hechos e hipótesis, ya sean patrones, reglas, grupos, funciones, modelos, secuencias, relaciones, correlaciones, etc. Esta tecnología, si bien reciente en su concepción, puede erigirse en una herramienta de obligada aceptación y uso en el competitivo entorno empresarial actual. Su utilización ya está extendida con cierta amplitud permitiendo, con éxito, la segmentación de mercados, la detección de hábitos de compra puntuales y secuenciales en el tiempo, la prevención del fraude y la morosidad en operaciones de crédito, la anticipación y gestión de problemas de salud de los empleados, etc. En definitiva, se trata de ejemplos donde cabe apreciar una cualidad común (entre otras más): la posibilidad de anticiparse a las variaciones del entorno, lo que facilitará darles una mejor y más rápida respuesta.

Por tanto, la MD se puede interpretar como el proceso, máximamente optimizado, intermedio entre la información y la toma de decisiones asociada a la misma. La aplicación ideal de la MD se llevaría a cabo sobre las bases de datos corporativas, que pueden ser un AD, o sobre otras específicas de propósito departamental (o Data Marts), contemplando elementos como los siguientes:

- Agentes inteligentes: Se encargan de analizar la información para detectar patrones y relaciones, ya de forma automática ya interactuando con el analista. Las técnicas que utilizan les permiten identificar grupos, comportamientos, reglas cuyo descubrimiento habría supuesto un enorme esfuerzo de trabajo metódico.
- Detección de alarmas: Consiste en la ejecución periódica o permanente de ciertos agentes para detectar acciones o situaciones susceptibles de desencadenar una acción extraordinaria o fuera del ciclo ordinario, pudiéndose activar en tiempo real o pueden detectarse y almacenarse para posterior análisis y tratamiento.
- Análisis multidimensional: Se basa en la estructuración y presentación de la información bajo aquellas perspectivas, ejes o dimensiones de interés.
- Consultas e informes: Las plataformas suelen incorporar herramientas de consulta con interfaces gráficos muy avanzados, intuitivos y fáciles de usar, cierto grado de análisis multidimensional y agentes inteligentes.
- Tratamiento de datos: Los datos suelen estar almacenados en los formatos más adecuados para su gestión por parte de los sistemas existentes, pero pueden no ser los más adecuados para su procesamiento por parte de la MD, de ahí que muchos desarrollos de MD incorporan módulos de tratamiento de datos con el objeto de simplificar al máximo las interfaces de datos e información.

En consecuencia, cabe poner de manifiesto que las aplicaciones de MD extraen conocimiento escondido, patrones de comportamiento no explícitos, relaciones ocultas o información predictiva del almacén, sin necesidad de preguntas o peticiones específicas sino utilizando distintas técnicas, tales como algoritmos matemáticos, métodos estadísticos, modelos lógicos borrosos, algoritmos genéticos, inducciones de reglas, sistemas expertos y sistemas basados en el conocimiento y redes neuronales.

De hecho, aún cuando el estado del arte en lo que se refiere a soluciones operativas de MD es todavía incipiente, ya es posible plantear, a modo de resumen, alguna de las herramientas que utilizan o suelen utilizar las soluciones de MD, a saber:

1. **Agrupamiento.** Esta herramienta posibilita la identificación de tipologías o grupos en los cuales los elementos guardan similitud entre sí y se diferencian de los de otros grupos. A través de este tipo de herramientas se pueden segmentar, por ejemplo, el colectivo de clientes, el conjunto de valores e índices financieros, el conjunto de empleados, etc., lo que facilita el tratamiento particularizado de cada uno de estos grupos.

2. **Asociación.** Permite establecer las posibles relaciones entre acciones o sucesos aparentemente independientes. Así, se puede reconocer cómo la ocurrencia de un determinado suceso puede inducir la aparición de otro u otros. Este tipo de herramientas son particularmente útiles, por ejemplo, para comprender los hábitos de compra de los clientes y para la concepción de ofertas, de ventas cruzadas y del "merchandising".
3. **Secuenciamiento.** Es un concepto similar al anterior, pero incluyendo también el factor tiempo, esto es, permite reconocer el tiempo que transcurre o suele transcurrir entre el suceso inductor y los sucesos inducidos.
4. **Reconocimiento de patrones.** Permite la asociación de una señal o información de entrada con aquella o aquellas con las que guarda mayor similitud, y que están ya catalogadas en el sistema.
5. **Clasificación.** Agrupa a todas las herramientas que permiten asignar a un elemento la pertenencia a un determinado grupo o clase, a través de la dependencia de la pertenencia a cada clase en los valores de una serie de atributos o variables. Se establece un perfil característico de cada clase y su expresión, en términos de un algoritmo o reglas, en función de las distintas variables y también el grado de discriminación o influencia de estas últimas. De esta forma, es posible clasificar un nuevo elemento una vez conocidos los valores de las variables presentes en él.
6. **Simulación.** Estas herramientas permiten comparar la situación actual de la empresa y su posible evolución futura con otras posibles situaciones, para a partir de ahí decidir si se deben o no modificar las políticas de actuación en determinados ámbitos de la misma.
7. **Optimización.** Trata de resolver el problema de la minimización o maximización de una función que depende de una serie de variables, encontrando los valores de éstas que satisfacen la condición de máximo, típicamente beneficios, o mínimo, típicamente costes. En general, suelen tener lugar unas restricciones, que hacen que no todas las posibles soluciones sean aceptables, de modo que el universo de búsqueda se reduce a aquellas soluciones que satisfagan las restricciones.
8. **Previsión.** Permite establecer el comportamiento futuro más probable de una variable o una serie de variables a partir de la evolución pasada y presente de esas variables o de otras de las cuales dependan.

3. HEURÍSTICAS BASADAS EN LA NATURALEZA

En la última década se ha producido, en prácticamente todos los campos del quehacer científico, una importante transformación conceptual y metodológica ligada íntimamente al estudio de los llamados fenómenos no-lineales cuyo análisis se engloba, parcialmente, dentro de las llamadas ciencias de la complejidad o de los sistemas complejos. Como parte de esta nueva visión, se ha puesto en evidencia que diversas propiedades espacio-temporales de los sistemas complejos surgen espontáneamente a partir de interacciones entre los elementos constituyentes, en escalas de tiempo y longitud considerablemente mayores que las escalas en las que ocurren dichas interacciones. Estas propiedades, denominadas "propiedades emergentes", han comenzado a ser estudiadas con una familia nueva de herramientas y conceptos originadas en la interacción interdisciplinaria de varios campos de la ciencia, desde la física, la biología, la química, la economía, la sociología, etc.

Los sistemas complejos están formados por un conjunto grande de componentes individuales que interactúan entre sí y que pueden modificar sus estados internos como producto de tales interacciones. Tales sistemas pueden ser estructuralmente simples, aunque tal simplicidad no impide que exhiban comportamientos dinámicos diversos y no triviales. Asimismo, los sistemas complejos pueden situarse en regímenes críticos caracterizados por la presencia de fluctuaciones espaciales y temporales en todas las escalas posibles. Esta situación crítica puede alcanzarse espontáneamente y sin la intervención de factores o fuerzas externas al sistema, estableciendo entonces lo que se denomina "proceso autoorganizado".

El proceso de interacciones puede generar comportamientos colectivos y globales, es decir, conductas que no están definidas en los elementos individuales, pero que emergen como un proceso colectivo y que no pueden ser reducidas ni explicadas tomando aisladamente a los elementos constituyentes.

En la naturaleza existe un elevado número de ejemplos de sistemas complejos que van desde las reacciones químicas autocatalíticas, hasta los procesos sociales y culturales. La naturaleza posee una fuerte tendencia a estructurarse en forma de entes discretos excitables que interactúan y que se organizan en niveles jerárquicos de creciente complejidad; por ello, los sistemas complejos no son de ninguna manera casos raros ni curiosidades sino que dominan la estructura y función del universo. Tales sistemas constituyen y se manifiestan en la inmensa mayoría de los fenómenos observables. Sin embargo, y aquí radica una de sus propiedades más inte-

resantes, la abundancia y diversidad de los sistemas complejos (sean de tipo físicos, químicos, biológicos, sociales, etc.) no implica una innumerable e inclasificable diversidad de conductas dinámicas diferentes. Todo lo contrario, los sistemas complejos poseen propiedades genéricas, independientemente de los detalles específicos de cada sistema o de la base material del mismo.

Uno de los más importantes y prometedores campos de investigación han sido las denominadas “Heurísticas basadas en la Naturaleza” (HBNs), un área cuya denominación surge debido a que simulan los mecanismos de la naturaleza para plantear la resolución de problemas, esto es, se basan en emplear analogías con sistemas naturales o sociales, al objeto de diseñar heurísticos no determinísticos de “búsqueda”, “aprendizaje”, “comportamiento”, etc., al considerar la naturaleza como fuente de ideas para el desarrollo técnico y científico.

De acuerdo con lo anterior, a continuación se propondrá una taxonomía que pudiera presentar alguna utilidad para entender mejor las direcciones que siguen las actuales líneas de investigación en este campo. Conviene, no obstante, poner de manifiesto previamente que en la definición de este tipo de algoritmos, la metáfora “natural” hace referencia a los sistemas derivados de las ciencias físicas, biológicas y sociales. Por su parte, las heurísticas se obtienen: realizando un cierto número repetido de experimentos, utilizando uno o más “agentes” (neuronas, cromosomas, hormigas, etc.), operando (en el caso de múltiples agentes) con un mecanismo de competición-cooperación e incluyendo procedimientos de automodificación de los parámetros de la heurística o de la representación del problema.

Una aproximación a la evolución natural permite establecer los mecanismos básicos de la misma, a saber: la selección que premia a los individuos más fuertes y penaliza a los más débiles, y la mutación, que introduce elementos aleatorios y permite el nacimiento de nuevos individuos. En las HBNs se plantea una situación similar: la selección es la idea básica para la optimización mientras que la mutación es la idea básica para la búsqueda no determinística.

Entre las características principales de las HBNs cabe distinguir las siguientes: (i) modelan (de forma aproximada) un fenómeno existente en la naturaleza; (ii) son no determinísticas; y (iii) son adaptativas, esto es, la capacidad del sistema para utilizar realimentación de la información a través de la modificación de sus parámetros y su modelo interno.

Las características anteriores facilitan un “comportamiento razonable” para el sistema, que se podría definir como “inteligente” (capacidad para resolver problemas difíciles) que, en otros términos, significa la producción de buenas soluciones para un problema de optimización combinatorial. Todas las aproximaciones para resolver este tipo de problemas de manera exacta están basados en una enumeración implícita de las soluciones factibles y, por tanto, requieren, en el peor de los casos, un número exponencial de pasos en aras de facilitar la exploración del espacio de soluciones y, consecuentemente, para ganar seguridad en la optimalidad de la solución encontrada.

Entre las principales características que se persiguen equilibrar en la construcción de algoritmos heurísticos resaltan las dos siguientes: (i) el grado de explotación, es decir, la cantidad de esfuerzo empleado por la búsqueda local en la región actual del espacio de búsqueda (si una región es prometedora, la búsqueda se hace más a fondo); y (ii) el grado de exploración, esto es, la cantidad de esfuerzo gastado por la búsqueda en regiones distantes del espacio (en ocasiones se elige una solución en una región lejana y se acepta una solución peor para tener la posibilidad de descubrir nuevas soluciones mejores).

Las dos posibilidades anteriores presentan un conflicto en la construcción de este tipo de algoritmos, de forma que es preciso tratar de establecer un equilibrio entre ambas el cual se debe ajustar en cada algoritmo. Asimismo, la construcción de algoritmos heurísticos debe estar presidida por un equilibrio entre el esfuerzo requerido o eficiencia (entendido como número de iteraciones) y la eficacia (valor final de la solución).

3.1. Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNAs), se comportan imitando el funcionamiento del cerebro humano, basado en el aprendizaje a través de la experiencia, con la consiguiente extracción de conocimiento a partir de la misma, con lo cual las RNAs no solucionan problemas a través de una secuencia de pasos sino que, de

forma análoga al cerebro, utilizan la combinación de una gran cantidad de elementos simples de proceso (neuronas) interconectados entre sí (sinapsis) que operan en paralelo.

La teoría y modelado de RNAs se basa en la estructura y funcionamiento de los sistemas nerviosos biológicos; donde las neuronas (cuya etimología griega es “célula nerviosa”), como se muestra en la Figura 9, constituyen la unidad fundamental del sistema nervioso de los seres vivos, y particularmente del cerebro. Considerando su tamaño microscópico resulta absolutamente sorprendente su capacidad como procesador de señales eléctricas y su complejidad bioquímica.

Una neurona es una célula viva y, como tal, contiene los mismos elementos que forman parte de todas las células biológicas, si bien incluye además elementos característicos que la diferencian. En general, una neurona consta de un cuerpo celular más o menos esférico, de 5 a 10 micras de diámetro, del que salen una rama principal, el axón, y varias ramas más cortas, llamadas dendritas. A su vez, el axón puede producir ramas en torno a su punto de arranque, y con frecuencia se ramifica extensamente cerca de su extremo.

Una de las características que diferencian a las neuronas del resto de las células vivas es su capacidad de comunicarse. Así, en general, las dendritas reciben señales de entrada (inputs) procedentes de otras neuronas o bien de estímulos nerviosos exteriores. Estos inputs pasan al cuerpo celular, que los combina e integra (a través de funciones no lineales), emitiendo señales de salida a través del axón; dicho axón se encuentra conectado a través de sus ramificaciones a las dendritas de otras neuronas, que reciben estas señales y, siguiendo el proceso descrito, las combinan para producir nuevas salidas propagadas a través de sus axones a nuevas neuronas, etc.

La conexión entre el axón de una neurona y las dendritas de otra se llama sinapsis y determina la fuerza y tipo de relación entre las mismas. En general, como se muestra en la Figura 1, se estima que una neurona recibe información de miles de otras neuronas, transmitiendo dicha información, previamente procesada, a miles de neuronas más.

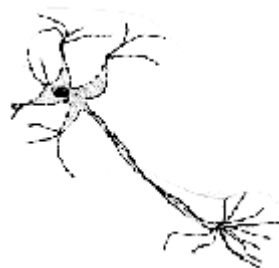


Figura 1

El hecho de que unas determinadas sinapsis sean excitadoras o inhibitoras y que tengan una mayor o menor importancia con respecto a la entrada total a la neurona, es establecido biológicamente a través del aprendizaje. Así, a medida que se va adquiriendo conocimiento las conexiones interneuronales (sinapsis) van modificando sus características, adaptándose con el fin de provocar respuestas satisfactorias ante la presencia de estímulos.

Por otro lado, una Red Neuronal Artificial (RNA) puede considerarse como un modelo matemático de las "teóricas" actividades mentales y cerebrales. Estos sistemas explotan el procesamiento local en paralelo y las propiedades de la representación distribuida, aspectos que al parecer existen en el cerebro biológico.

En este sentido, una RNA puede concebirse como una colección de procesadores elementales (neuronas), conectadas a otras neuronas o entradas externas y con una salida que permite propagar las señales por múltiples caminos. Cada procesador pondera las entradas que recibe, y estos pesos pueden ser modificados con el fin de conseguir el objetivo previsto, con lo cual una RNA puede "aprender" de sus propios errores. Por tanto, se trata de un Proceso Inductivo a partir de un conjunto de ejemplos de lo que se quiere aprender, frente al Proceso Deductivo, propio de los Sistemas Expertos.

Las RNAs imitan el funcionamiento del cerebro humano, basado en el aprendizaje a través de la experiencia, con la consiguiente extracción de conocimiento a partir de la misma; de ahí que no solucionan problemas a través de una secuencia de pasos sino que, mimetizando el comportamiento del cerebro, utilizan la combinación de una gran cantidad de elementos simples de proceso (neuronas) interconectados entre sí (sinapsis) que operan en paralelo. Así, puede definirse una RNA como: "un sistema o conjunto de procesadores elementales interconectados, no lineal ni estacionario, que realiza al menos alguna de las siguientes funciones: Aprendizaje, Memorización, Generalización o Abstracción de características esenciales".

Los modelos de RNAs combinan modelos matemáticos de las células nerviosas y modelos de arquitecturas que describen las interconexiones que existen entre estas células. Al igual que en el sistema nervioso, una red neuronal se configura a base de interconectar un conjunto de neuronas de forma que las salidas de unas constituyen las entradas de otras.

En este sentido, se supone que la información que recibe una neurona artificial, procedente de todas aquellas otras a las que se encuentra conectada, puede identificarse por un número que mide la frecuencia de la excitación del procesador. De esta forma, la integración de todas las excitaciones se reduce a un proceso de suma de los diferentes valores, afectados de unos coeficientes de ponderación, denominados pesos sinápticos. El resultado de esta suma representa la actividad interna de la célula, sirviendo de base para el cálculo de la señal que posteriormente será transmitida a posteriores neuronas. El esquema general de comportamiento de una neurona artificial aparece recogido en la Figura 2.

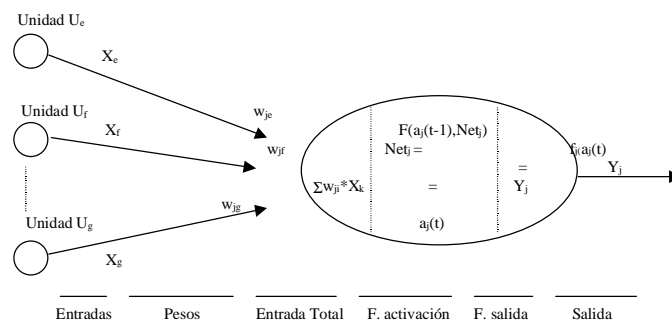


Figura 2

Los pesos sinápticos miden la intensidad de la sinapsis o conexión entre las neuronas receptoras y transmisoras y, al igual que en el caso de las neuronas biológicas, su función básica consiste en el almacenamiento de información. Así, puede decirse que cada una de las neuronas componentes de la red almacena una cantidad muy pequeña de información, la denominada memoria de trabajo a corto plazo, mientras que la memoria a largo se acumula en los pesos sinápticos, ya que éstos son los que determinan el carácter excitador o inhibitorio de las entradas y de las diferentes relaciones neuronales, así como su importancia relativa.

De acuerdo con lo anterior, el aprendizaje tiene lugar mediante la conveniente modificación de las conexiones neurales. De esta forma, puede definirse dicho aprendizaje como el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el proceso de aprendizaje se reducen a la destrucción, modificación y creación de conexiones entre las neuronas, entendiéndose por destrucción el proceso por el cual un peso pasa a tener valor 0 y por construcción el paso del valor 0 a un valor distinto de cero, si bien en algunos modelos de redes la construcción hace referencia también a la introducción de nuevas neuronas en la red, con los consiguientes pesos asociados y la destrucción a la salida de neuronas de la red, con la desaparición de los pesos a ellas vinculados.

Durante el proceso de aprendizaje los pesos de las conexiones de la red sufren modificaciones, por lo que se puede afirmar que el aprendizaje ha terminado cuando los valores de los pesos permanecen estables; en la práctica, en la mayoría de las redes, esto ocurre cuando el error acometido por la red es inferior a un determinado límite previamente fijado (5%, 3%, etc.) o bien cuando el número de épocas, esto es, las vueltas o veces

en las que los ejemplos de entrenamiento son presentados a la red para que ésta “aprenda”, supera un determinado valor.

Por lo que respecta al modo concreto en que se efectúa tal aprendizaje, el mismo depende del tipo de utilización que se pretenda implementar en la red, si bien en cualquier caso el aspecto esencial es que el ajuste se hace teniendo presentes de forma exclusiva los datos del problema que se quiere resolver, esto es, sin llevar a cabo clasificaciones previas de los mismos ni la formulación de hipótesis de partida que presupongan un modelo de cómo se relacionan. De esta forma, tal como se muestra en la Figura 3 es posible distinguir básicamente dos tipos de aprendizaje: (i) el aprendizaje supervisado, en el que los outputs relacionados con los ejemplos de aprendizaje son conocidos a priori; y (ii) el aprendizaje no supervisado, donde no se conocen las correspondientes salidas del sistema.

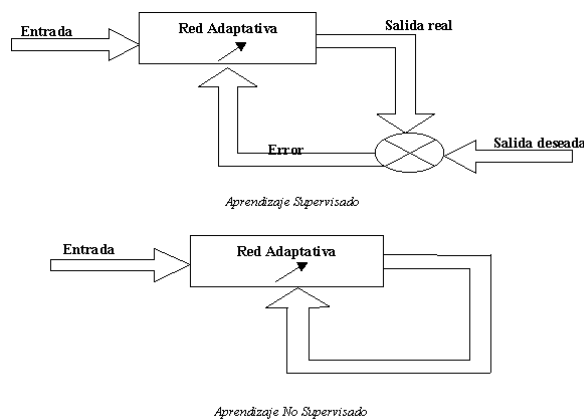


Figura 3

3.2. Algoritmos Genéticos

Los Algoritmos Genéticos (AGs) son métodos basados en el proceso genético de los organismos vivos, en el que los individuos de una población compiten entre sí, de forma que aquellos individuos con mayor éxito para sobrevivir tienen mayor probabilidad de generar un gran número de descendientes mientras que los pocos dotados darán lugar a un número de descendientes menor. En definitiva, esto significa que los genes de los individuos mejor adaptados se propagan en las generaciones siguientes hacia un número de individuos creciente, provocando que los individuos posteriores (“hijos”) tengan una adaptación mayor que la de sus progenitores (“padres”). De esta forma, las especies evolucionan logrando unas características cada vez mejor adaptadas al entorno en que se desarrollan.

La Evolución es un proceso no-dirigido, esto es, no existen pruebas científicas de que el proceso evolutivo esté encaminado a la consecución de un objetivo final. Más bien, puede verse como un proceso reactivo, en virtud del cual los organismos cambian en respuesta a las variaciones del entorno. En este sentido, cabe recordar que los principios básicos de la evolución de los seres vivos surgen a mediados del s. XIX, los cuales se asientan en la Teoría de la Selección Natural de Darwin en 1859 y los trabajos sobre Herencia Genética de Mendel en 1865.

Los AGs fueron introducidos por Holland (1975) como algoritmos de búsqueda que utilizan la metáfora de las poblaciones de genes. En la comunidad de los AGs, un problema de optimización se traslada al problema de encontrar la mejor adaptación de un individuo (llamado cromosoma) dentro de una población. La adaptación se mide por una función de fitness o función de adaptación (bondad), que está relacionada con la función objetivo del problema que se desea resolver. El AG opera sobre la población modificando sus componentes. Las modificaciones ocurren de acuerdo a unas reglas genéticas implementadas a través de los operadores genéticos (cruce y mutación). Los individuos son el equivalente a las soluciones y una población es un conjunto de N individuos. Cada individuo consiste en una colección (generalmente una cadena) de elementos atómicos llamados genes. Cada gen puede tomar valores dentro de un conjunto predefinido. De esta forma, el comportamiento natural traducido en los AGs presenta las siguientes fases:

En primer término, se genera una población de individuos cada uno de los cuales representa una solución factible al problema que se desea resolver.

En función de la bondad de la solución que representa cada individuo, se le asigna una valoración que en definitiva establece el grado de efectividad del individuo para competir por unos determinados recursos. Cuanto mayor sea la adaptación de un individuo mayor será la probabilidad de que dicho individuo sea seleccionado para reproducirse y, en consecuencia, para que su material genético se propague en sucesivas generaciones.

El proceso de reproducción se realiza cruzando el material genético del individuo con el de otro individuo seleccionado de igual forma, generando una nueva población que reemplaza a la anterior con la ventaja de que esta población contiene una mayor proporción de buenas características que la anterior.

A través de sucesivas generaciones las nuevas poblaciones estarán mejor adaptadas que las poblaciones de las que provienen sin más que favorecer el cruce de los individuos con mejores características, ya que de esta forma se van explorando las áreas más prometedoras del espacio de búsqueda. Si el AG está bien diseñado, en un tiempo razonable la población convergerá hacia una solución óptima del problema.

En consecuencia, las fases de un AG se pueden resumir en el esquema que muestra la Figura 4.

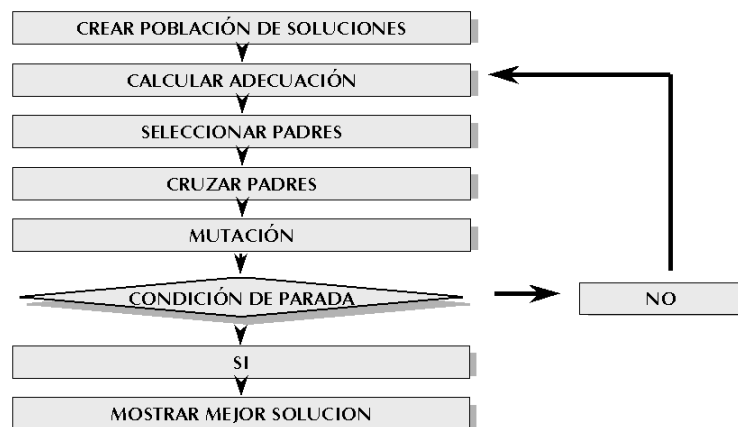


Figura 4

De acuerdo con lo anterior, para poder aplicar un AG se requieren los siguientes componentes básicos: (i) una representación de las soluciones potenciales del problema; (ii) una forma de crear una población inicial de posibles soluciones; (iii) una función de evaluación que juegue el papel del entorno o ambiente, clasificando las soluciones en términos de su “aptitud”; (iv) operadores genéticos que alteren la composición de los hijos que se producirán para las siguientes generaciones; y (v) valores para los diferentes parámetros que utiliza el AG, entre los que cabe destacar el tamaño de la población, la probabilidad de cruce, la probabilidad de mutación, el número de generaciones, etc.

El algoritmo genético básico representa los individuos (posibles soluciones del problema) mediante un conjunto de parámetros a los que se denominan genes, los cuales se agrupan constituyendo una cadena de valores que se define como cromosomas. Si bien existen distintas posibilidades de llevar a cabo la codificación de los individuos, las cuales se comentarán con posterioridad, en general la codificación más utilizada se basa en el alfabeto binario $\{0, 1\}$.

El conjunto de parámetros representando un cromosoma particular se denomina fenotipo, el cual contiene la información requerida para construir un organismo, el cual se denomina genotipo. Por tanto, la adecuación de un individuo a un problema depende de la evaluación del genotipo, la cual puede inferirse a partir del fenotipo, es decir, puede ser computada a partir del cromosoma mediante la función de evaluación.

Por su parte, la función de evaluación debe ser específica para cada problema y servirá para asignar un valor a cada cromosoma en particular, el cual reflejará el nivel de adaptación al problema del individuo representado

por el cromosoma. Los principales problemas derivados de la formulación del modelo son la convergencia prematura y la finalización lenta.

Para constituir la siguiente generación de individuos se realiza un proceso de reproducción que consiste, en primer término, en seleccionar los individuos para someterlos al operador de cruce y constituir los descendientes, para posteriormente realizar estos últimos un proceso de mutación.

La selección de individuos que van a actuar como “padres” se realiza normalmente al azar, utilizando un procedimiento que favorezca a los individuos mejor adaptados, para lo cual se asigna a cada individuo una probabilidad que es proporcional a su función de adaptación. A este procedimiento se le conoce con la denominación de “ruleta sesgada” cuyo objetivo es que los individuos mejor adaptados sean seleccionados, incluso varias veces por generación, mientras que aquellos individuos con una adaptación pequeña se crucen en muy pocas ocasiones. No obstante, existen distintas posibilidades para realizar la selección de los individuos que se considerarán “padres” para la siguiente generación que se analizarán más adelante.

La principal ventaja de los AGs es que se trata de técnicas robustas que se pueden utilizar con éxito en una gran variedad de problemas incluso en casos en que otras técnicas plantean dificultades. Por otra parte, si bien no se puede garantizar que un AG proporcione una solución óptima, empíricamente se ha demostrado que encuentran soluciones aceptables en un tiempo competitivo con el resto de algoritmos de optimización combinatoria.

3.3. Algoritmos Basados en Hormigas

Los algoritmos basados en hormigas fueron propuestos por Coloni, Dorigo y Maniezzo (1991) como un sistema de múltiples agentes para resolver problemas combinatoriales difíciles, como el viajante de comercio (Stützle y Dorigo, 1992a) y el problema de la asignación cuadrática. En la actualidad existen varias propuestas que extienden y aplican los algoritmos basados en hormigas, y que se recogen bajo la denominación de “Algoritmos de Optimización mediante Colonias de Hormigas” o “Algoritmos ACO” (Stützle y Dorigo, 1992b; Dorigo y Di Caro, 1999). Estos algoritmos están inspirados en la observación de colonias de hormigas reales: se trata de insectos sociales que viven en colonias y que tienen un comportamiento dirigido al desarrollo de la colonia como un todo más que a un desarrollo individual.

Una característica interesante del comportamiento de las colonias de hormigas es cómo pueden encontrar los caminos más cortos entre el hormiguero y la comida. Estudios entomológicos descubrieron que esta capacidad es el resultado de la interacción debido a la comunicación química entre las hormigas, a través de una sustancia llamada feromona, y un fenómeno emergente causado por la presencia simultánea de muchas hormigas. En este sentido, y a efectos ilustrativos, cabe considerar el caso de la Figura 5 que sigue a continuación, donde se muestra unas hormigas reales que se mueven en una línea recta que conecta una fuente de comida con su hormiguero.



Figura 5

Pero sucede que también son capaces de adaptarse a los cambios en el entorno, por ejemplo buscando un nuevo camino más corto cuando debido a un obstáculo el camino antiguo resulta más largo. Dicha adaptación se produce por el hecho de que las hormigas depositan una cierta cantidad de feromona mientras caminan y cada hormiga prefiere probabilísticamente seguir una dirección rica en feromona que otra más pobre en dicha sustancia.

Este elemental comportamiento de las hormigas puede ser utilizado para explicar como buscan el camino más corto que reconecte la línea rota que, en no pocas ocasiones, acontece ante un inesperado obstáculo que interrumpe el camino inicial, tal como se muestra en la Figura 6.

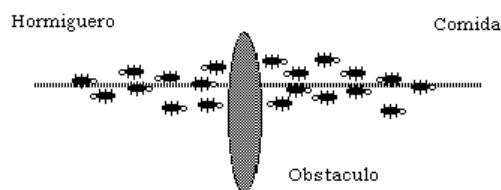


Figura 6

De hecho, cuando aparece un impedimento, las hormigas que están justamente al frente no pueden continuar siguiendo la pista de feromona y entonces deben escoger entre torcer a la derecha o a la izquierda. En esta situación, como no tienen indicaciones sobre la mejor opción, cabría esperar que eligieran de forma aleatoria: la mitad de las hormigas escogerán torcer a la derecha y la otra mitad a la izquierda, tal como se muestra en la Figura 7.

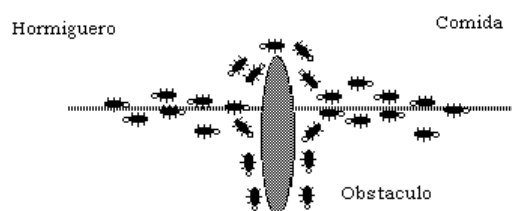


Figura 7

No obstante, resulta interesante comprobar que las hormigas que han escogido, por suerte, el camino más corto del obstáculo, pueden reconstituir más rápidamente la pista de feromona comparadas con las que escogieron el camino más largo. De ahí que el camino más corto pueda recibir una mayor cantidad de feromona en la misma cantidad de tiempo, siendo ésta la causa de que un mayor número de hormigas seleccionen el camino más corto.

Debido a este proceso (autocatalítico) de retroalimentación positiva, muy pronto todas las hormigas escogen el camino más corto, tal como se muestra en la Figura 8.

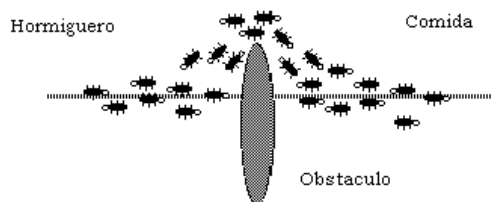


Figura 8

A partir de este momento, las nuevas hormigas preferirán con mayor probabilidad elegir el camino de arriba, ya que en el punto de decisión percibirán una mayor cantidad de feromona proveniente de este camino.

El aspecto más interesante de este proceso autocatalítico radica en que la búsqueda del camino más corto alrededor del obstáculo permite la observación de una propiedad emergente entre el contorno o forma del obstáculo y el comportamiento distribuido de las hormigas: aunque las hormigas se mueven aproximadamente a la misma velocidad y depositan una tasa de feromona similar, el tardar más cuando se desplazan por el contorno más largo que por el más corto propicia que la preferencia por seguir pistas de feromona con mayor acumulación sirva para escoger este último camino. Este procedimiento se puede generalizar para recorridos con varios puntos o nodos, de forma que cada vez que una hormiga llega a una intersección, decide el camino a seguir de un modo probabilístico. Las hormigas eligen con mayor probabilidad los caminos con un alto rastro de feromona. De esta forma, las bifurcaciones más prometedoras (más cercanas a la comida) van acumulando feromona en tanto son recorridas por más hormigas, mientras que las menos prometedoras pierden

feromona por evaporación al ser visitadas por menos hormigas cada vez. Por tanto, la acción continuada de la colonia da lugar a un rastro de feromona que permite a las hormigas encontrar un camino cada vez más corto desde el hormiguero a la comida, ya que la cantidad de feromona depositada en un arco es inversamente proporcional a su longitud.

En relación a los distintos desarrollos operativos en este ámbito cabe destacar especialmente los Sistemas de Hormigas (SHs) por ser los primeros algoritmos ACO que modelan el comportamiento de las colonias de hormigas reales. En su aplicación a un problema real es necesario que el mismo pueda ser representado en forma de grafo, de forma que la Figura 7, representativa de una colonia de hormigas reales, puede ser representada como se refleja en la Figura 9.

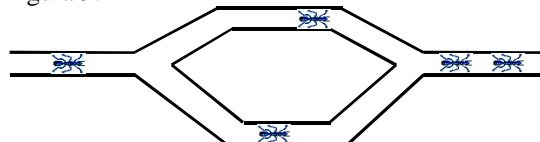


Figura 9

Un ejemplo ilustrativo de este funcionamiento puede ser el que muestra el grafo de la Figura 10 (adaptado de Dorigo, Maniezzo y Coloni, 1996):

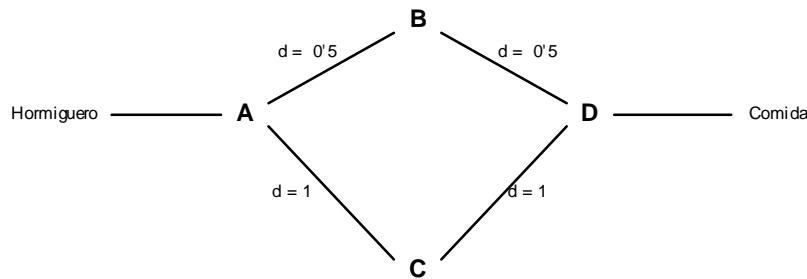


Figura 10

Los arcos (AB) y (BD) suponen la mitad de distancia que la existente entre los arcos (AC) y (CD), de forma que si se supone que las hormigas caminan a la misma velocidad, tardarán la mitad de tiempo en su recorrido. Si se supone que en el momento de tiempo t , 30 hormigas salen del hormiguero hacia la comida, y otras 30 hormigas parten en la dirección contraria, al llegar al punto de intersección correspondiente (A o D), elegirán el camino a seguir de forma aleatoria (por ejemplo, con una probabilidad del 50%), tal como recoge la Figura 11.

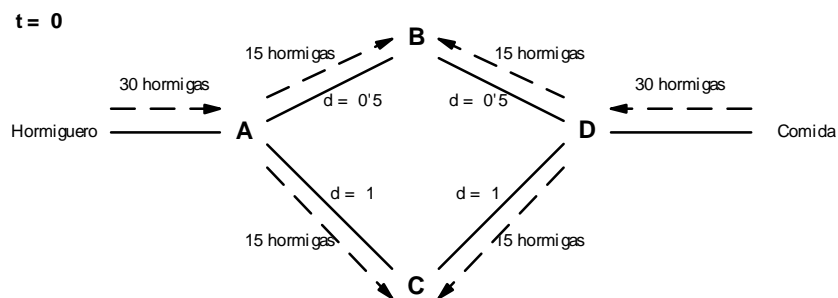


Figura 11

Si se supone que cada hormiga deja un rastro de feromona igual a la unidad y que, para mayor simplicidad, la feromona se evapora de forma total en cada unidad de tiempo, en el momento $t+1$, los arcos (AB y DB) habrán sido recorridos por 30 hormigas (15 en cada sentido), mientras que por los arcos (AC y DC) sólo habrán pasado 15 hormigas. De esta forma, la feromona recogida en cada arco, al considerar que cada hormiga aporta una unidad de feromona, será de 30 unidades para los arcos más cortos y 15 unidades para los arcos que re-

presentan recorridos más largos. En el momento $t+1$, salen nuevamente 30 hormigas del hormiguero y de la comida. Dado que el rastro de feromona es mayor en los arcos superiores, la probabilidad de elección de los mismos será superior, de forma que se producirá una situación como la que refleja la Figura 12.

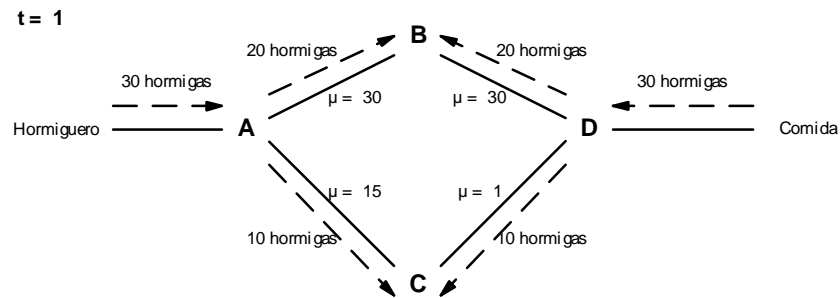


Figura 12

En consecuencia, la distancia entre dos arcos establece una medida del coste dada entre dos nodos r y s , $\delta(r, s)$, de tal forma que cada arco (r, s) tiene asimismo asociada una medida de deseabilidad, $\tau(r, s)$ denominada feromona, que se actualiza por las hormigas artificiales (o, simplemente, hormigas) en tiempo de ejecución.

El problema puede ser representado como un grafo con pesos, en el que cada arco del grafo contiene dos tipos de información distintas y con funciones diferentes, a saber:

Información heurística, que representa una medida del coste del arco, dependiente del caso concreto, que se calcula antes de comenzar el algoritmo y no se modifica durante la ejecución del algoritmo (en el ejemplo, la distancia entre los arcos).

Información memorística, que proporciona una medida de la “deseabilidad” del arco, representada por la cantidad de feromona depositada en él y modificada durante la ejecución del algoritmo en función del número de hormigas que recorrieron el mismo en el pasado (en el ejemplo, la feromona). En los sistemas basados en colonias de hormigas, el aporte de feromona depende también de la bondad de las soluciones que generaron las hormigas que recorrieron cada arco.

De esta forma, se puede definir una hormiga artificial como un agente que: (i) recuerda los nodos que ha recorrido, utilizando para ello una lista tabú de nodos visitados (L); (ii) tras cada iteración, esta lista contiene la solución construida por la hormiga; (iii) en cada paso, elige hacia qué nodo moverse (qué arista seguir) de entre los alcanzables desde el actual r que no hayan sido visitados aún ($J(r) = \{u \mid \exists(r, u) \text{ y } u \notin L\}$), según una regla probabilística de transición; (iv) una vez construida su solución, deja un rastro de feromona τ_{ij} (en una cantidad que depende de la bondad de la misma) en cada arco por el que ha pasado y vacía L ; y (v) opcionalmente, puede también depositar feromona en cada arco que recorre mientras construye la solución.



Figura 13

El funcionamiento básico del SHs, tal como se muestra en la Figura 13, es el siguiente: en cada iteración, una población de H hormigas construye progresivamente, según una regla de transición de estados que depende de la información existente, distintos recorridos por el grafo (soluciones al problema). Una vez evaluadas éstas, los arcos asociados a las soluciones más prometedoras son reforzados por un aporte adicional de feromona, mientras que la contenida en los demás arcos del grafo es evaporada.

De acuerdo con lo anterior, para exponer el funcionamiento del SH es preciso determinar la forma de establecer la regla de transición de estados y la regla de actualización de la feromona:

La regla de transición de estados utilizada por el sistema de hormigas, denominada regla proporcional-aleatoria, define una distribución de probabilidad para el hecho de que una hormiga k en un nodo r elija para moverse el nodo s y que viene dada por la siguiente expresión:

$$p_k(r, s) = \begin{cases} \frac{[\tau(r, s)]^\alpha \cdot [\eta(r, s)]^\beta}{\sum_{u \in J_k(r)} [\tau(r, u)]^\alpha \cdot [\eta(r, u)]^\beta}, & \text{si } s \in J_k(r) \\ 0, & \text{en otro caso} \end{cases}$$

donde:

$\tau(r, s)$ nivel de feromona del arco (r, s)

$\eta(r, s)$ información heurística que, en el caso de venir establecida en términos de coste se determinará como $\eta = 1/\delta$, es decir, el inverso del coste del arco $(\delta(r, s))$

$J_k(r, s)$ conjunto de los nodos alcanzables desde el nodo r no visitados aún por la hormiga k (para hacer la solución factible)

α y β parámetros que determinan la importancia relativa de la feromona en relación con la información heurística. En general $\beta > 0$ mientras que el valor de α suele ser 1, razón por la cual en ocasiones este parámetro es obviado.

En la regla de transición de estados se multiplica la feromona del arco (r, s) por el correspondiente valor heurístico $\eta(r, s)$ con la finalidad de favorecer la elección de los arcos más prometedores (de menor coste) y con mayor cantidad de feromona.

Una vez que cada hormiga ha generado su solución, la regla de actualización global de feromona modifica el nivel de feromona de cada arco del grafo de dos formas principales, a saber: (i) evaporando feromona en los arcos que no fueron visitados por ninguna hormiga en la iteración actual (arcos poco prometedores) y (ii) añadiendo feromona en los visitados en función de la bondad de la solución que generó la hormiga que los visitó (arcos prometedores).

La expresión de la regla de actualización global de feromona es:

$$\tau(r, s) = \underbrace{(1 - \rho) \cdot \tau(r, s)}_{\text{Evaporación}} + \underbrace{\sum_{k=1}^H \Delta \tau_k(r, s)}_{\text{Aporte}}$$

donde:

ρ parámetro de evaporación de feromona ($\rho \in [0, 1]$)

H número de hormigas

$\Delta \tau_k(r, s) = \begin{cases} f(S_k), & \text{si la hormiga } k \text{ ha visitado el arco } (r, s) \\ 0, & \text{en otro caso} \end{cases}$

$f(S_k)$ cantidad de feromona proporcional a la bondad de la solución generada por la hormiga k

De esta forma, si la medida de la bondad de la solución se establece en función del coste de la solución obtenida por cada hormiga (L_k), entonces:

$\Delta \tau_k(r, s) = \begin{cases} 1/L_k, & \text{si } (r, s) \in \text{solución construida por la hormiga } k \\ 0, & \text{en otro caso} \end{cases}$

Como puede observarse, todas las hormigas aportan feromona a los arcos de sus soluciones en función de la bondad de éstas. Aquellos arcos que no forman parte de ninguna solución sufren la evaporación de un $(1-\rho)$ por ciento de la feromona de que disponían.

Este proceso de actualización de la feromona tiene el propósito de asignar más cantidad de feromona a las soluciones de mayor bondad (menor coste). En este sentido, es similar a un esquema de aprendizaje por refuerzo en el cual las mejores soluciones obtienen un mayor refuerzo (como ocurre, por ejemplo, en los algoritmos genéticos cuando se utiliza como mecanismo de selección el método proporcional a la bondad de las soluciones). La fórmula de actualización de la feromona controla el cambio de la cantidad de ésta sobre los arcos, tanto para añadir más cantidad de feromona en los arcos visitados como para evaporarla de los arcos no prometedores.

La acción de situar feromona sobre los arcos simula el papel de una memoria distribuida a largo plazo: esta memoria no se almacena localmente dentro de las hormigas individuales, sino que está distribuida sobre los arcos del grafo. Esto permite una forma indirecta de comunicación entre ellas.

CONCLUSIONES

El propósito de este trabajo documento fue presentar alguna de las tecnologías emergentes, las denominadas heurísticas bio-inspiradas, que pueden servir de apoyo en las actividades de minería de datos de utilidad en el marketing público, considerando como marketing público el conjunto de actividades cuyo objetivo es el diseño, implementación y control de programas destinados a satisfacer las necesidades de los usuarios o consumidores de los servicios brindados por una entidad pública, mediante el adecuado diseño del servicio, la distribución, la promoción, el personal, de la presencia física, de los procesos y eventualmente, de los precios.

Pues, aunque existen diferencias importantes entre marketing privado y marketing público, especialmente en cuanto a los objetivos de la entidad, la forma de establecer los objetivos de marketing, las estructuras de mercado en que se desenvuelven las entidades, el concepto de consumidor en uno y otro caso, el entorno legal, algunos de los criterios de segmentación, el grado de control sobre la determinación de precios y grado de uso de algunos de los métodos promocionales. Sin embargo, todo parece indicar son más las similitudes que las diferencias: en ambos casos es necesario lograr la satisfacción de los usuarios o consumidores del servicio mediante el uso de ciertas variables controlables.

Por consiguiente, parece evidente que el instrumental utilizado en el marketing privado es también utilizable en el marketing público.

BIBLIOGRAFÍA

- Coloni, A.; Dorigo, M. y Maniezzo, V. (1991): "Distributed Optimization by Ant Colonies". Incluido en Varela, F. y Bourgine, P. (eds.): "Proceedings of ECAL-91 - European Conference on Artificial Life". Elsevier, Paris, págs.134-142.
- Dorigo, M. y Di Caro, G. (1999): "The Ant Colony Optimization Meta-Heuristic". Incluido en Corne, D.; Dorigo, M. y Glover, F. (eds.): "New Ideas in Optimization". McGraw-Hill, New York, págs. 25-38.
- Dorigo, M.; Maniezzo, V. y Coloni, A. (1996): "The Ant System: Optimization by a Colony of Cooperating Agents". IEEE Transactions on Systems, Man, and Cybernetics-Part B, Vol.26, nº 2, págs. 29-41.
- Holland, J. (1975): "Adaptation in Natural and Artificial Systems". Univ. Michigan Press, Ann Arbor (MIT Press, 1992).
- Kohonen, T. (1989): "Self Organization and Adaptive Memory". Springer Verlag. Berlín.
- López González, E. y Flórez López, R. (2000): "Aplicación de dos modelos de redes neuronales artificiales para el análisis económico-financiero empresarial". Revista Europea de Dirección y Economía de la Empresa. Vol. 9, nº 2, págs. 141-165.
- López González, E. y Rodríguez Fernández, M. (2000): "Genetic Optimisation of a Fuzzy Distribution Model". International Journal of Physical Distribution and Logistics Management. Vol. 30, nº 7/8, págs. 681-696.
- López González, E.; Mendaña Cuervo, C. y Rodríguez Fernández, M. (1998a): "Fijación de estrategias del mix de promoción mediante un algoritmo genético borroso". XIX Congrès de la Association Française Comptabilité, Nantes, mayo, págs. 442-443.
- López González, E.; Rodríguez Fernández, M.; Mendaña Cuervo, C. y Flórez López, R. (1999): "Distribution Information Systems with Genetic Algorithm and Fuzzy Sets". 1999 EUSFLAT – ESTYLF Joint Conference, Universidad de las Islas Baleares, Palma de Mallorca, septiembre, págs-323-326.

- López González, E; Mendaña Cuervo, C. y Rodríguez Fernández, M. (1997b): "Global Sourcing: An Approach to Suppliers Selection Problems using Genetic Algorithm and Linguistic Valuations". VI International Conference of the European Association of Management and Business Economics, Chania (Grecia), septiembre, págs. 117-130.
- López González, E; Rodríguez Fernández, M. y Mendaña Cuervo, C. (2001): "The logistic Decision Making in Management Accounting with Genetic Algorithms and Fuzzy Sets". *Mathware & Soft Computing*, Vol. 7, págs. 1-15.
- Mendaña Cuervo, C. (2000): "Modelos de gestión basados en tecnologías bio-inspiradas para el desarrollo de nuevos productos". Tesis Doctoral. Servicio de Publicaciones de la Universidad de León. León
- Parker, D. (1982): "Learning logic. Invention Report", Office of Technology Licensing, Universidad de Stanford. Stanford.
- Rumelhart, D. et al. (1986): "Learning representations by back-propagating errors", *Nature*, 323, pp. 533-536.
- Stützle, T. y Dorigo, M. (1992a): "ACO Algorithms for the Travelling Salesman Problem". Université Libre de Bruxelles, Bruxelles.
- Stützle, T. y Dorigo, M. (1992b): "ACO Algorithms for the Quadratic Assignment Problem". Université Libre de Bruxelles, Bruxelles.
- Werbos, P. (1974): "Beyond Regression: New tools for prediction and analysis in the behavioral sciences". Tesis Doctoral. Universidad de Harvard. Harvard.