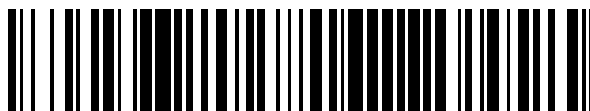


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 611 177**

51 Int. Cl.:

**G06N 3/08** (2006.01)

**G06N 3/04** (2006.01)

**G06F 17/28** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.03.2012 E 12159672 (0)**

97 Fecha y número de publicación de la concesión europea: **16.11.2016 EP 2639749**

54 Título: **Procedimientos, aparatos y productos para el procesamiento semántico de texto**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**05.05.2017**

73 Titular/es:  
**CORTICAL.IO GMBH (100.0%)**  
**Mariahilfer Straße 4/11**  
**1070 Vienna, AT**

72 Inventor/es:  
**WEBBER, FRANCISCO EDUARDO DE SOUSA**

74 Agente/Representante:  
**ZEA CHECA, Bernabé**

**ES 2 611 177 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Procedimientos, aparatos y productos para el procesamiento semántico de texto

5 **Campo de la invención**

La presente invención se refiere a un procedimiento de entrenamiento de una red neuronal, en particular para el procesamiento semántico, clasificación y predicción de texto. La invención se refiere, además, a medios legibles por ordenador y máquinas de clasificación, predicción y traducción basadas en redes neuronales.

10

Antecedentes de la invención

En el contexto de la presente descripción, el término "*red neuronal*" designa una red neuronal artificial implementada en ordenador. En Bishop C. M., "*Neuronal Networks for Pattern Recognition*", Oxford University Press, Nueva York, 1995/2010; o Rey, G. D., Wender K. F., "*Neuronale Netze*", 2ª edición, Hans Huber, Hofgrete AG, Berna, 2011, por ejemplo, se da una visión general de la teoría, tipos y detalles de implementación de redes neuronales.

15

La presente invención se ocupa, en particular, del procesamiento semántico de texto a través de redes neuronales, es decir, analizando el significado de un texto, centrándose en la relación entre sus palabras y lo que representan en el mundo real y en su contexto. A continuación, "palabras" (fichas) de un texto comprenden tanto palabras en la terminología habitual del idioma, así como cualquier unidad de un idioma que pueda combinarse para formar un texto, tal como símbolos y signos. A partir de estas palabras, se prescinde de un grupo de palabras con demasiada ubicuidad como "el/la", "él", "a/en" etc. que tienen poca relevancia semántica que dejar lo que se denomina "palabras clave" de un texto.

20

25

Las aplicaciones de un procesamiento de texto semántico son generalizadas y abarcan, por ejemplo, la clasificación de texto en determinadas palabras clave para clasificar la relevancia, archivar, minería de datos y fines de recuperación de información. Comprender el significado de palabras clave en un texto y predecir otras palabras clave "con significado" que se produzcan en el texto es útil, por ejemplo, para la expansión de consultas semánticas en motores de búsqueda. Por último, pero no menos importante, el procesamiento de texto semántico mejora la calidad de las traducciones automáticas mediante la resolución de ambigüedades de un texto de origen al considerar sus palabras en un contexto semántico más amplio.

30

Los procedimientos de procesamiento de texto semántico existentes hasta la fecha, en particular para la expansión de consultas en motores de búsqueda, funcionan con grandes índices estadísticos para palabras clave, su lema (raíces léxicas) y relaciones estadísticas entre las palabras clave para construir grandes archivos de sinónimos, estadísticas y diccionarios para el análisis relacional. Los procedimientos estadísticos, sin embargo, están limitados en profundidad de análisis semántico cuando se consideran secuencias de palabras más largas y más complejas.

35

Por otra parte, las redes neuronales se utilizan principalmente para el reconocimiento de patrones en datos complejos y diversos, tales como el reconocimiento de objetos en imágenes o el reconocimiento de señales en datos de voz, música o de medición. Las redes neuronales tienen que ser "entrenadas" correctamente con cantidades masivas de datos de entrenamiento con el fin de poder cumplir con su tarea de reconocimiento cuando se alimenta con muestras "vivas" a analizar. El entrenamiento de una red neuronal es equivalente a configurar sus conexiones internas y pesos entre sus nodos de red ("neuronas"). El resultado del entrenamiento es una configuración específica de conexiones ponderadas normalmente dentro de la red neuronal.

40

45

El entrenamiento de una red neuronal es una tarea compleja por sí misma e implica el establecimiento de una multitud de parámetros, por ejemplo, con algoritmos iterativos o de adaptación. Por lo tanto, los algoritmos de entrenamiento para redes neuronales pueden ser considerados como un medio técnico para construir una red neuronal para una aplicación específica.

50

Para reducir la dimensionalidad de los vectores para entrenar una red neuronal, es conocido, de Ampazis N. y otros, "*A Latent Semantic Indexig Approach to Self-Organizing Maps of Document Collections*", Neuronal Processing Letters, Kluwer Academix Publishers, 2004, págs. 1 - 17, procesar previamente vectores de entrada basados en palabras, truncándolos mediante una Descomposición de Valores Singulares y reduciendo su número, categorizándolos a través de una primera red neuronal, y entrenando después una segunda red neuronal con vectores basados en categorías en lugar de vectores basados en palabras.

55

Mientras que las redes neuronales están actualmente en un uso extendido para el reconocimiento de patrones en grandes cantidades de datos numéricos, su aplicación en el procesamiento de texto está actualmente limitada por la forma en que puede presentarse un texto a una red neuronal en una forma legible por una máquina.

60

Descripción de la invención

Un objetivo de la invención es mejorar la interfaz entre el texto, por una parte, y las redes neuronales por otra, con el fin de aprovechar mejor la potencia de análisis de redes neuronales para el procesamiento de texto semántico.

5 En un primer aspecto de la invención, se dispone un procedimiento implementado por ordenador para el entrenamiento de una red neuronal, que comprende:

10 entrenar una primera red neuronal de tipo de mapa de auto-organización con un primer grupo de primeros documentos de texto que contienen cada uno una o más palabras clave en un contexto semántico para asignar cada documento a un punto en el mapa de auto-organización mediante agrupación semántica;  
15 determinar, para cada palabra clave que tiene lugar en el primer grupo, todos los puntos en el mapa de auto-organización a los cuales se asignan primeros documentos que contienen dicha palabra clave como un patrón y guardar dicho patrón para dicha palabra clave en un diccionario de patrones;  
20 formar por lo menos una secuencia de palabras clave a partir de un segundo grupo de documentos de texto conteniendo cada uno una o más palabras clave en un contexto semántico;  
traducir dicha por lo menos una secuencia de palabras clave en por lo menos una secuencia de patrones utilizando dicho diccionario de patrones; y  
entrenar una segunda red neuronal con dicha por lo menos una secuencia de patrones.

La segunda red neuronal entrenada con el innovador procedimiento está configurada y lista para ser utilizada en una variedad de aplicaciones, incluyendo las siguientes aplicaciones:

- 25 i) procesamiento de texto que contiene por lo menos una palabra clave, que comprende:  
traducir dicha por lo menos una palabra clave en por lo menos un patrón por medio del diccionario de patrones, enviar dicho por lo menos un patrón como patrón de entrada a dicha segunda red neuronal entrenada,  
30 obtener por lo menos un patrón de salida de dicha segunda red neuronal entrenada, y traducir dicho por lo menos un patrón de salida en por lo menos una palabra clave por medio del diccionario de patrones;
- 35 ii) clasificación semántica de texto, si se utiliza una segunda red neuronal de tipo jerárquico, en el que dicho por lo menos un patrón de entrada se envía a por lo menos una capa inferior de la jerarquía y dicho por lo menos un patrón de salida se obtiene a partir de por lo menos una capa superior de la jerarquía; y
- 40 iii) predicción semántica de texto, si se utiliza una segunda red neuronal de tipo jerárquico, en el que dicho por lo menos un patrón de entrada se envía a por lo menos una capa superior de la jerarquía y dicho por lo menos un patrón de salida se obtiene a partir de por lo menos una capa inferior de la jerarquía.

45 En otro aspecto, la invención presenta un procedimiento para generar un diccionario legible por ordenador para traducir texto en una forma legible por una red neuronal, que comprende:

50 entrenar una red neuronal de tipo de mapa de auto-organización con documentos de texto, conteniendo cada uno una o más palabras clave en un contexto semántico para asignar cada documento de texto a un punto en el mapa de auto-organización mediante agrupación semántica;  
determinar, para cada palabra clave que tiene lugar en el primer grupo, todos los puntos en el mapa de auto-organización al cual se asignan documentos de texto que contienen dicha palabra clave, como un patrón de puntos asociados a dicha palabra clave; y  
guardar todas las palabras clave y patrones asociados como un diccionario legible por ordenador.

55 La invención también presenta un diccionario legible por ordenador de este tipo que se realiza en un medio legible por ordenador.

Otros aspectos de la invención son:

60 - una máquina de clasificación, que comprende una red neuronal de tipo de memoria temporal jerárquica que ha sido entrenada como dicha segunda red neuronal con un segundo procedimiento de acuerdo con el primer aspecto de la invención;

- una máquina de predicción, que comprende una red neuronal de tipo de memoria temporal jerárquica que ha sido entrenada como dicha segunda red neuronal con un procedimiento de acuerdo con el primer aspecto de la invención;

5 - una máquina de traducción, que comprende dicha máquina de clasificación, cuya red neuronal ha sido entrenada utilizando primeros y segundos documentos de texto en un primer idioma, y una máquina de predicción, cuya red neuronal ha sido entrenada utilizando primeros y segundos documentos de texto en un segundo idioma, en el que los nodos de la red neuronal de la máquina de clasificación están conectados a los nodos de la red neuronal de la máquina de predicción.

10 En todos los aspectos, la invención combina tres tecnologías diferentes de una manera completamente novedosa, es decir, mapas auto-organizados (SOMs), la indexación inversa de palabras clave en un SOM, y una red neuronal objetivo expuesta al texto traducido en un flujo de patrones.

15 Uno de los principios de la invención es la generación de un nuevo tipo de un diccionario "palabra clave vs. patrón" (abreviado: el "diccionario de patrones") que contiene una asociación entre una palabra clave y un patrón de dos (o más) dimensiones. Este modelo representa la semántica de la palabra clave en el contexto del primer grupo de documentos. Seleccionando una colección de contextos semánticos como primer grupo de documentos, por ejemplo, artículos de una enciclopedia, tal como se describirá más adelante, cada patrón refleja el contexto semántico y, por lo tanto, el significado de una palabra clave.

20 Los patrones son generados por una red neuronal SOM, en particular un "mapa de auto-organización de Kohonen" ("mapa de características de Kohonen"). Para más detalles de SOMs véase, por ejemplo de Kohonen, T. "*The Self-Organizing Map*", Proceedings of the IEEE, 78 (9), 1464-1480, 1990; Kohonen, T., Somervuo, P., "*Self-Organizing Maps of Symbol Strings*", Neurocomputing, 21 (1-3), 19-30, 1998; Kaski, S., Honkela, T., Lagus, K., Kohonen, T., "*Websom-Self-Organizing Maps of Document Collections*", Neurocomputing, 21(1-3), 101-117, 1998; Merkl, D., "*Text Classification with Self-Organizing Maps: Some Lessons Learned*", Neurocomputing, 21 (1-3), 61-77, 1998; Vesanto, J., Alhoniemi, E., "*Clustering of the Self-Organizing Map*", IEEE Transactions on Neuronal Networks, 11 (3), 586-600, 2000; Pözlbauer G., M. Dittenbach, Rauber A., "*Advanced Visualization of Self-Organizing Maps with Vector Fields*", IEEE Transactions on Neuronal Networks 19, 911-922, 2006.

30 Los patrones generados por SOM se utilizan posteriormente para traducir secuencias de palabras clave de un segundo grupo (entrenamiento) de documentos de texto en secuencias de patrones para enviarse a la segunda de red neuronal (objetivo) para el reconocimiento de patrones. El reconocimiento de patrones es una de las principales competencias de las redes neuronales. Puesto que cada modelo representa un significado intrínseco de una palabra clave, y una secuencia de patrones representa un significado contextual de palabras clave, la semántica de las palabras clave en el segundo grupo de documentos es analizada por la red neuronal objetivo en referencia y ante el origen del significado intrínseco de las palabras clave en el contexto del primer grupo de documentos. Como resultado, la red neuronal objetivo puede analizar de manera eficiente y de manera significativa la semántica de un texto.

40 Los procedimientos y aparatos de la invención son adecuados para el entrenamiento de todo tipo de redes neuronales objetivo. Una aplicación preferida es la formación de redes neuronales que son jerárquicas y - por lo menos parcialmente - recurrentes, en particular, redes neuronales de tipo de marco de predicción de memoria (MPF) o memoria temporal jerárquica (HTM). Para detalles de teoría e implementación de MPFS y EMATs véase, por ejemplo, Hawkins, J., George, D., Niemasik, J., "*Sequence Memory for Prediction, Inference and Behaviour*", Philosophical Transactions of the Royal Society of London, Serie B, Biological Sciences, 364(1521), 1203-9, 2009; Starzyk, J. A., He, H., "*Spatio-Temporal Memories for Machine Learning: A Long-Term Memory Organization*", IEEE Transactions on Neural Networks, 20(5), 768-80, 2009; Numenta, Inc., "*Hierarchical Temporal Memory Including HTM Cortical Learning Algorithms*", Whitepaper of Numenta, Inc., Versión 0.2.1, 12 de septiembre de 2011; Rodríguez A., Whitson J., Granger R., "*Derivation and Analysis of Basic Computational Operations of Thalamocortical Circuits*", Journal of Cognitive Neuroscience, 16:5, 856-877, 2004; Rodríguez, R. J., Cannady, J. A., "*Towards a Hierarchical Temporal Memory Based Self-Managed Dynamic Trust Replication Mechanism in Cognitive Mobile Ad-hoc Networks*", Actas de la 10ª Conferencia Internacional WSEAS sobre inteligencia artificial, ingeniería del conocimiento y bases de datos, 2011; así como las patentes (solicitudes) Nos. US 2007/0276774 A1, US 2008/0059389 A1, US 7 739 208 B2, US 7 937 342 B2, US 2011/0225108 A1, US 8 037 010 B2 y US 8 103 603 B2.

60 Las redes neuronales MPF y HTM guardan representaciones jerárquicas y secuenciadas en el tiempo de flujos de patrones de entrada y son particularmente adecuadas para comprender semántica de texto en el tiempo y jerárquica. Sus nodos (neuronas) en diferentes capas jerárquicas representan *per se* abstracciones jerárquicas (clases) de palabras clave; la clasificación (abstracción) es un principio de funcionamiento intrínseco de este tipo de redes cuando la entrada se envía de abajo a arriba de la jerarquía, y la predicción (detalle) es un principio de funcionamiento intrínseco cuando la entrada se envía de arriba a abajo de la jerarquía.

En otro aspecto de la invención el concepto de nodos que representan clases enteras (abstracciones, categorías) de palabras clave se utiliza para construir una máquina de traducción como una máquina de predicción asignada a salidas de nodos de una máquina de clasificación.

5 De acuerdo con otro aspecto de la invención, varios segundos documentos pueden utilizarse y traducirse en flujos de patrones de entrenamiento para entrenar la segunda red neuronal en un grupo específico de segundos documentos.

10 En algunas realizaciones de la invención, los segundos documentos están ordenados por complejidad ascendente y, cuando se entrena la segunda red neuronal, las secuencias independientes de patrones se envían a la segunda red neuronal en el orden de clasificación de los segundos documentos a partir de los cuales se ha formado y traducido cada uno. Esto da lugar a un entrenamiento más rápido de la segunda red neuronal.

15 En algunos otros aspectos de la invención, la complejidad de un segundo documento se determina en base a uno o más de: el número de palabras clave diferentes en ese segundo documento, la longitud media de una frase en ese segundo documento, y la frecuencia de una o más palabras clave del primer grupo en ese segundo documento.

#### Breve descripción de los dibujos

20 La invención se describe más en detalle en referencia a los dibujos adjuntos, en los cuales:

La figura 1 es un diagrama de flujo de resumen general del procedimiento de la invención, incluyendo diagramas de bloques de una primera y una segunda red neuronal, un diccionario de patrones, así como máquinas de clasificación, predicción y traducción de acuerdo con la invención;

25 La figura 2 es un diagrama de flujo de la etapa de procesamiento de vectores para el primer documento establecido como vector de entrada a la primera red neuronal de la figura 1;

La figura 3 es un mapa de auto-organización (SOM) de ejemplo creado como salida de la primera red neuronal de la figura 1;

30 La figura 4 es un diagrama de flujo de la etapa de indexación inversa, recibiendo entradas desde la etapa de procesamiento de vectores y el SOM, para crear el diccionario de patrones de la figura 1;

La figura 5 muestra representaciones de SOM inversamente indexados con patrones de ejemplo para dos palabras clave diferentes dentro del SOM;

La figura 6 muestra ejemplos de algunos patrones predeterminados para las palabras vacías (no palabras clave);

35 La figura 7 es un diagrama de flujo de la etapa de extracción de secuencias de palabras clave para el segundo grupo de segundos documentos en la figura 1;

La figura 8 muestra el resultado de una etapa de selección de documentos opcional para los segundos documentos del segundo grupo;

La figura 9 es un diagrama de flujo de las etapas de la traducción de una secuencia de palabras clave en una secuencia de patrones de la figura 1; y

40 La figura 10 muestra una estructura de nodos jerárquica de ejemplo de un MPF utilizado como segunda red neuronal de la figura 1.

#### Descripción detallada de la invención

45 En una perspectiva general, la figura 1 muestra un procedimiento de procesamiento de texto semántico y un sistema 1 que utiliza un primer grupo 2 de primeros documentos de texto 3 para formar una primera red neuronal 4. La primera red neuronal 4 es un mapa de tipo auto-organización (SOM) y crea un mapa de auto-organización (SOM) 5. A partir del SOM 5 se crean uno patrones 6 representativos de palabras clave 7 que se producen en el primer grupo de documentos 2 a través de una etapa de indexación inversa 8 y ponen en un diccionario de patrones 9.

50 El diccionario de patrones 9 se utiliza en una etapa de traducción 10 para traducir secuencias de palabras clave 11 extraídas de un segundo grupo 12 de segundos documentos 13 en un patrón de secuencias 14. Con el patrón de secuencias 14 se entrena una segunda red neuronal 15. La segunda red neuronal 15 es preferiblemente (aunque no necesariamente) de tipo marco de predicción de memoria (MPF) o memoria temporal jerárquica (HTM). La segunda red neuronal entrenada 15 puede utilizarse entonces para clasificar semánticamente texto traducido con un diccionario de patrones 9, véase camino 16, o bien para predecir semánticamente texto traducido con un diccionario de patrones 9, véase camino 17. Otra aplicación opcional de la segunda red neuronal entrenada 15 es una asignación jerárquica, véase caminos 18, a una tercera red neuronal opcional 19, que tiene una configuración similar a la segunda red neuronal 15, pero ha sido entrenada en un idioma diferente al de la segunda red neuronal 15; unas asignaciones de nodos 18 representan entonces coincidencias semánticas entre nodos semánticos 15' de la primera red de idiomas 15 y nodos semánticos 19' de la segunda red de 19 idiomas.

Se describen ahora en detalle los procesos y funciones de los componentes mostrados en la figura 1 con referencia a las figuras 2-10.

5 La figura 2 muestra una etapa de preprocesamiento y vectorización 20 para indexar y vectorizar el primer grupo 2 de primeros documentos 3. En la etapa 20 a partir del primer grupo 2 se produce una secuencia de vectores de entrada 21, un vector 21 para cada primer documento 3, como grupo o matriz (tabla) de vectores de entrenamiento de entrada 22 aplicado a la capa de entrada 23 de la primera red neuronal (SOM) 4. Tal como es conocido para el experto en la materia, la red neuronal SOM 4 normalmente comprende sólo dos capas, una capa de entrada 23 y una capa de salida 24 de las neuronas (nodos), interconectadas por unas conexiones 25 cuyos pesos pueden ser representados por una matriz de ponderación. Las redes neuronales SOM pueden ser entrenadas con algoritmos de aprendizaje sin supervisión en el que los pesos de la matriz de ponderación se auto-adaptan a los vectores de entrada, para asignar específicamente nodos de la capa de entrada 23 a nodos de la capa de salida 24, teniendo en cuenta la relación espacial de los nodos de la capa de salida 24 en un mapa de dos dimensiones (o más) 5. Esto produce mapas 5 que agrupan vectores de entrada 21 respecto a su similitud, produciendo regiones 26 en el mapa 5 con vectores de entrada muy similares 21. Para detalles de las redes neuronales SOM, véase las referencias bibliográficas citadas anteriormente.

20 El primer grupo 2 y los primeros documentos 3 en el mismo se seleccionan en tal número y granularidad, por ejemplo, longitud de los documentos individuales 3, de manera que cada uno de los documentos 3 contiene un grupo de, por ejemplo, 1 a 10, 1 a 20, 1 a 100, 1 a 1000 o más, preferiblemente aproximadamente 250 a 500, palabras clave 7 en un contexto semántico. Un primer documento 3 puede contener - además de las palabras clave 7 - palabras de poca relevancia semántica (tales como artículos "un/una", "el/la" etc.) que normalmente se denominan palabras vacías, aquí no palabras clave.

25 El número de documentos 3 en el grupo 2 se selecciona para obtener un corpus representativo de contextos semánticos para las palabras clave 7, por ejemplo, miles o millones de documentos 3. En una realización de ejemplo, se utilizan aproximadamente 1.000.000 documentos 3, cada uno comprendiendo aproximadamente entre 250 y 500 palabras clave 7, como primer grupo de documentos 2.

30 La longitud (recuento de palabras clave) de los documentos 3 debe ser bastante constante en todo el grupo 2, las palabras clave 7 deben estar distribuidas de manera uniforme y escasamente en los documentos 3 en el grupo 2, y cada documento 3 debe contener una buena diversidad de palabras clave 7.

35 Las palabras clave 7 también pueden ser raíces (lema) de palabras, de modo que, por ejemplo, para formas de singular y plural (gato/gatos) o de diferentes formas verbales (ir/yendo) sólo se tiene en cuenta una palabra clave 7. Las palabras clave 7, por lo tanto, pueden ser formas de palabras específicas y/o raíces de las palabras. Después de eliminar palabras incapaces de construir palabras clave importantes, tales como palabras vacías, cada documento 3 puede considerarse una "bolsa de palabras" de palabras clave 7.

40 En una realización práctica, puede generarse un primer grupo adecuado 2, por ejemplo, a partir de artículos de una enciclopedia, tales como artículos de Wikipedia® obtenidos bajo la "*Creative Commons Attribution Licence*" o la "*GNU Free Documentation Licence*" del proyecto Wikipedia®. Tales artículos enciclopédicos, o entradas, respectivamente, pueden analizarse de acuerdo con capítulos, párrafos etc. en documentos 3 de longitud bastante uniforme, de modo que cada documento 3 contiene palabras clave 7 en un contexto semántico, es decir, con significado.

45 Para generar los vectores 21, se genera un índice de todas las palabras clave 7 que existen en todo el grupo 2 y se extiende horizontalmente como encabezamiento de columna 27 de la matriz (tabla) 22. A la inversa, identificaciones de documentos ("id") de todos los documentos 3 en todo el grupo 2 se extienden verticalmente como encabezamiento de fila 28 de la matriz 22. Después, para cada vez que tiene lugar una palabra clave específica 7 en un documento específico 3, se pone una marca o binario "1" en la respectiva celda de la matriz 22. Por lo tanto, en la matriz 22 una fila horizontal representa un vector "palabra clave-ocurrencia" normalizado 21 para un documento 3, en el que un binario "1" en una posición de palabra clave específica (posición de la columna) indica que esta palabra clave 7 está contenida en la "bolsa de palabras" de este documento 3; y un binario "0" indica la ausencia de esta palabra clave 7 en este documento 3. O, a la inversa, cada columna de la matriz 22 muestra para una palabra clave específica 7 todos aquellos documentos 3 marcados con un binario "1" que contiene esa palabra clave 7.

50 Los vectores de entrada 21, es decir, filas de la matriz 22 que representan los documentos 3 y su contenido de palabras clave, se suministran entonces sucesivamente a la capa de entrada 23 de la red neuronal SOM 4 para entrenarla. Esto significa que, si se utiliza un primer grupo 2 de, por ejemplo, 1.000.000 primeros documentos 3, se envía una sesión de entrenamiento de 1.000.000 entradas de vectores a la primera red neuronal 4.

- Como resultado de esta sesión de entrenamiento, la capa de salida 24 de la red neuronal SOM 4 ha producido un mapa 5 en el cual se han asignados documentos 3 (vectores 21) a puntos individuales ("píxeles")  $X_i/Y_j$  del mapa 5, agrupados por semejanza. La figura 3 muestra un ejemplo de un mapa 5. A cada punto del mapa  $X_1/Y_1, X_2/Y_2, \dots, X_i/Y_j, \dots$ , se le ha asignado cero, uno o más documentos 3 con su bolsa de palabras clave 7. Los documentos 3 (vectores 21) se identifican en el mapa 5, por ejemplo, por su identificador de documento a partir del encabezado de fila 28. Mediante ese proceso de agrupación de SOM, diferentes documentos 3, que contienen palabras clave 7 muy similares, por ejemplo, que coinciden en un 80% o un 90% de sus palabras clave, se asignan en estrecha relación espacial entre sí, formando así "regiones" semánticas 26<sub>a</sub>, 26<sub>b</sub>, 26<sub>c</sub>, 26<sub>d</sub>, etc. en el mapa 5.
- A continuación, en la etapa de indexación inversa 8 de la figura 4, en la base de la matriz 22 para una determinada palabra clave 7 a partir de un índice de palabras clave 27 se identifican todos aquellos documentos 3 que contienen la palabra clave 7. Esto puede realizarse fácilmente, por ejemplo, recuperando todo binario "1" en la columna específica de la palabra clave determinada 7 en la matriz 22 y buscando el identificador del documento 3 listado en el encabezado de la fila 28.
- Para aquellos documentos 3 que se ha comprobado que contienen esa palabra clave determinada 7, todos los puntos del mapa  $X_i/Y_j$  que referencian esa identificación del documento específica se determinan a partir del mapa 5. Este grupo  $\{X_i/Y_j\}$  de puntos del mapa representa el patrón 6. El patrón 6 es representativo de los contextos semánticos en que existió esa palabra clave determinada 7 en el primer grupo 2: La distribución espacial (es decir, de dos o más dimensiones) de los puntos  $X_i/Y_j$  en el patrón 6 refleja aquellas regiones semánticas específicas 26<sub>a</sub>, 26<sub>b</sub>, ... en cuyo contexto existió la palabra clave 7 en el primer grupo 2.
- El patrón 6 puede codificarse como un mapa binario 31, véase la figura 4, y también considerarse como una "huella dactilar" o "huella de pisada" binaria del significado semántico de una palabra clave 7 en una colección de documentos, tales como el primer grupo 2. Si el primer grupo 2 cubre una amplia variedad de textos con significado en un idioma específico, el patrón 6 es de elevada relevancia semántica de la palabra clave 7.
- La resolución espacial del patrón 6 puede ser igual o menor que la resolución espacial de la red neuronal SOM 4 y/o el mapa 5. La resolución espacial de este último puede seleccionarse de acuerdo con el rendimiento de análisis requerido: Por ejemplo, un mapa 5 puede estar compuesto de millones de puntos de mapa  $X_i/Y_j$ , por ejemplo 1000 x 1000 puntos, y el patrón 6 pueden tener la misma resolución para una elevada precisión, o una resolución menor para menores requerimientos de memoria.
- La figura 5 muestra un ejemplo de dos patrones diferentes 6 (representados como puntos negros) cubriendo el mapa 5 para la facilidad de comprensión. En este ejemplo, las regiones 26<sub>a</sub>, 26<sub>b</sub>, 26<sub>c</sub>, 26<sub>d</sub> han sido etiquetadas manualmente con clases semánticas tales como "depredador", "felinos", "mi mascota" y "canis". Esto es sólo a modo de ejemplo; hay que señalar que no es necesario un etiquetado para el correcto funcionamiento de los presentes procedimientos, procesos y algoritmos que sólo requieren la distribución de SOM espacial de los puntos del mapa  $X_i/Y_j$ .
- En la representación de la izquierda de la figura 5, todos los documentos 3 en que existía la palabra clave "gato" han sido marcados con un punto. En la representación de la derecha de la figura 5, todos los documentos que contienen la palabra clave 3 "perro" se han marcado con un punto. Puede apreciarse fácilmente que los documentos "gato" caen principalmente, o se agrupan, en regiones 26<sub>b</sub> ("mi mascota") y 26<sub>d</sub> ("felinos"), mientras que los documentos "perro" 3 se agrupan principalmente en regiones 26<sub>b</sub> ("mi mascota") y 26<sub>c</sub> ("canis").
- Volviendo a la figura 1, para cada palabra clave 7 que existe en el primer grupo 2, el respectivo patrón 6 se guarda en el diccionario de patrones 9 en forma de asignación de dos vías, es decir, la asociación entre una palabra clave 7 y su patrón 6. El diccionario de patrones 9 constituye un primer producto intermedio del procedimiento y el sistema 1 de la figura 1. El diccionario de patrones 9 puede guardarse ("realizarse") en un medio legible por ordenador, por ejemplo, un portador de datos, tal como un disco duro, CD-ROM, DVD, chip de memoria, servidor de Internet, un almacenamiento en la nube en Internet, etc.
- Cabe señalar que la generación del diccionario de patrones 9 puede implicar el uso de una masiva potencia de procesamiento para la formación de la primera red neuronal 4 y el mapa de indexación inversa 5. Por lo tanto, el diccionario de patrones 9 preferiblemente se calcula previamente una vez y después puede utilizarse repetidamente en otras etapas y módulos de los procesos y las máquinas de la figura 1.
- En base a diferentes primeros grupos 2 de primeros documentos 3 que, por ejemplo, pueden seleccionarse específicos de la aplicación y/o y específicos del idioma, diferentes diccionarios de patrones 9 pueden calcularse previamente y distribuirse sobre medios legibles por un ordenador a aquellas entidades que realizan las etapas posteriores e implementan los módulos posteriores de los procesos y máquinas que se describirán ahora en detalle.

En estas etapas posteriores y módulos, la segunda red neuronal (objetivo) 15 es entrenada para un procesamiento de texto semántico en base al segundo grupo 12 de segundos documentos 13. Aunque el segundo grupo 12 podría ser idéntico al primer grupo 2, en la práctica, el segundo grupo 12 puede comprender un subgrupo del primer grupo 2 o, de hecho, segundos documentos específicos de la aplicación bastante diferentes 13. Por ejemplo, aunque el primer grupo 2 comprenda un gran número de documentos generales ("enciclopédicos") 3, el segundo grupo 12 puede ser un grupo de datos de usuario específicos de la aplicación de documentos de usuario 13 que, por ejemplo, requieran ser buscados por expansión (palabra clave) de consulta semántica, clasificados u ordenados por clasificación semántica, o traducidos por traducción semántica. El diccionario de patrones 9 refleja entonces el conocimiento semántico de fondo sobre los significados semánticos generales de las palabras clave 7, mientras que la segunda red neuronal 15 realiza un análisis en profundidad de un grupo de datos de usuario 12 de documentos de usuario 13.

Los documentos de usuario 13 pueden ser, por ejemplo, registros de bases de datos de productos, páginas web, documentos de patentes, registros médicos o todo tipo de colecciones de datos que serán analizados por la segunda red neuronal 15. Un requisito previo para el segundo grupo 12 es que haya sido escrito en el mismo idioma que el primer grupo 2, ya que, de lo contrario, el diccionario de patrones 9 podría no aplicarse de manera significativa al segundo grupo 12. Además, es preferible - aunque no es obligatorio - que palabras clave 7 que existen en los segundos documentos 13 del segundo grupo 12 estén comprendidas dentro de todo el grupo, es decir, el índice 27, de palabras clave 7 en el primer grupo 2, de modo que las palabras clave 7 del segundo grupo 12 se listen y puedan buscarse en el diccionario de patrones 9.

En el diccionario de patrones 9, palabras vacías o palabras clave pueden ignorarse o bien incorporarse como patrones simbólicos predeterminados o preconfigurados tales como los mostrados en la figura 6.

Para el entrenamiento de la segunda red neuronal 15, en una primera etapa 32 se extraen secuencias 11 de palabras clave 7 del segundo grupo 12. Las figuras 1, 7 y 8 muestran esta etapa de extracción en detalle. Básicamente sería suficiente si sólo un segundo documento 13 o unos pocos es (son) leído(s) secuencialmente, palabra a palabra, línea a línea, párrafo a párrafo, capítulo a capítulo, documento a documento, en una secuencia de lectura normal 33. Las palabras vacías o que no son palabras clave podrían omitirse (tratarse por separado, tal como se describe en la figura 6), y el resultado es una secuencia 11 de palabras clave 7. Preferiblemente, sin embargo, el segundo grupo 12 está dividido en una multitud de segundos documentos 13, y se genera una secuencia 11 de palabras clave 7 para un documento 13. Las secuencias 11 se utilizan entonces, - por ejemplo, en el orden de los documentos 13 a partir de los cuales se originan - como entrada de entrenamiento para la segunda red neuronal 15.

El entrenamiento de la segunda red neuronal 15 puede acelerarse si en la etapa de extracción 32 se realiza una clasificación opcional de los documentos 13 y/o secuencias 11. Para esta clasificación opcional, se calcula un "factor de complejidad" CompF en un proceso 34 para cada documento 13 del segundo grupo 12. El factor de complejidad CompF puede calcularse en base a uno o más de los siguientes parámetros de un documento 13:

- el número de palabras clave diferentes 7 en un documento 13;
- el recuento de palabras promedio de una frase o párrafo en un documento 13;
- la frecuencia, o diversidad, de una o más de las palabras clave 7, por ejemplo, de todas las palabras clave 7 del primer grupo 2, en un documento 13;
- la frecuencia de una o más de las palabras clave 7, por ejemplo, todas las palabras clave 7, de un documento 13 en el primer grupo entero 2 u otro corpus de texto representativo de lenguaje coloquial, por ejemplo, una colección de periódicos.

En la etapa de extracción 32 los documentos 13 pueden ordenarse (clasificarse) entonces de acuerdo con el factor de complejidad CompF ascendente, véase figura 8. De esta manera, la segunda red neuronal 15 recibe secuencias 11 de complejidad creciente, por ejemplo, primero se utilizan secuencias primitivas o simples 11 o secuencias 11 con una modesta diversidad de palabras clave 7, y después se utilizan secuencias 11 con estructuras semánticas y lingüísticas complicadas para entrenar la segunda red neuronal 15.

Antes de enviarse a la segunda red neuronal 15, las secuencias 11 de palabras clave 7 se traducen en una etapa de traducción 10 en base al diccionario de patrones 9. Cada palabra clave 7 en una secuencia 11 se busca en un diccionario de patrones 9, se recupera el patrón asociado 6, y los resultados son secuencias 14 de patrones 6, una secuencia de patrones 14 para cada documento 13. Cada secuencia de patrón 14 puede considerarse como una serie de tiempo o "clip de película" de patrones 6 que representan el contexto semántico de palabras clave 7 en un documento 13 en el contexto semántico global del primer grupo de documentos 2.

Cabe señalar que, en realizaciones simples, sería suficiente utilizar sólo una secuencia larga 14 de patrones 6 para entrenar la segunda red neuronal 15. Preferiblemente, se utiliza un gran número de secuencias de patrones 14 (una "secuencia de secuencias"), representando cada secuencia de patrones 14 un vector de entrenamiento con línea de



tiempo (matriz) para la segunda red neuronal 15. La figura 9 muestra un ejemplo de la etapa de traducción 10 que traduce una secuencia de palabras clave 11 en una secuencia de patrones 14.

5 En la etapa de entrenamiento (flecha 35 en la figura 1) la segunda red neuronal 15 recibe sucesivamente secuencias de patrones 14 para aprender los patrones 6 y sus secuencias en el tiempo. Tal como se ha descrito al principio, pueden utilizarse todo tipo de redes neuronales adaptadas para el procesamiento de grupos temporales de patrones, por ejemplo, redes neuronales que procesen patrones de alimentación de avance con ventanas deslizantes. Alternativamente y preferiblemente, pueden utilizarse redes neuronales recurrentes o por lo menos parcialmente recurrentes, con o sin bucles de retardo, para aprender y recordar secuencias temporales, por ejemplo, 10 redes neuronales asociativas por sí mismas o auto-asociativas.

15 En realizaciones ventajosas, la segunda red neuronal 15 también es jerárquica ya que capas superiores de la jerarquía comprenden un número de nodos (neuronas) menor que las capas inferiores de la jerarquía. La figura 10 muestra un ejemplo de dicha red jerárquica, en particular un marco de predicción de memoria (MPF), que también contiene conexiones de retroalimentación laterales (intra-capas, véase figura 1) y verticales (capa transversal) para el aprendizaje de secuencias temporales. Una forma preferida de tal arquitectura de MPF son redes neuronales de memoria de tipo temporal jerárquica (HTM). En los documentos anteriores se describe teoría y detalles de implementación de redes neuronales MPF y HTM.

20 Las redes HTM y MPF desarrollan - en la configuración entrenada - neuronas (nodos) dentro de la jerarquía, que representan abstracciones (clasificaciones) de patrones de activación de las neuronas (nodos) en capas inferiores de la jerarquía. Mediante el uso de conexiones intra-capas y capas transversales recurrentes entrenadas (retroalimentación), en particular entre nodos de estructuras de sub-capas "columnares", éstas pueden modelar el comportamiento temporal de flujos temporales enteros de patrones de activación. De esta manera, las redes HTM y 25 MPF pueden aprender, recordar y clasificar flujos de patrones y reconocer secuencias de patrones, así como predecir posibles secuencias de patrones futuros de secuencias de patrones anteriores.

30 Una vez que la red neuronal 15 ha sido entrenada con el patrón de secuencias 14, pueden aplicarse nuevos patrones 6 o nuevas secuencias de patrones 14 como nuevas entradas a una entrada de "clasificación" en niveles de menor jerarquía de la red 15, para obtener clasificaciones/abstracciones semánticas como patrones a partir de las salidas de los nodos en niveles jerárquicos superiores, véase ruta 16; o pueden introducirse nuevos patrones 6 o nuevas secuencias de patrones 14 en entradas de "predicción" en niveles jerárquicos superiores y pueden obtenerse patrones previstos (predicciones semánticas) en niveles inferiores de la jerarquía, véase ruta 17.

35 Tal como puede apreciarse en la figura 1, se utiliza un diccionario de patrones 9 en ambas rutas 16, 17 para traducir cualquier nueva secuencia de "consulta" de palabras clave 7 en una secuencia de "consulta" 14, y volver a traducir los patrones de salida de la red neuronal 15 en una clasificación "resultante" o palabras clave de predicción 7.

40 La ruta de clasificación 16, por lo tanto, puede utilizarse para clasificar un texto de consulta por la red neuronal entrenada 15 utilizando el diccionario de patrones 9 en las interfaces de entrada y salida de la red 15; y puede utilizarse una ruta de predicción 17 para predecir palabras clave a partir de un texto de consulta, por ejemplo, para "expandir" una frase de palabras clave de búsqueda a otras palabras clave (previstas) 7 que coincidan semánticamente con la frase de consulta (valor de referencia), utilizando un diccionario de patrones 9 en interfaces tanto de entrada como de salida de la red neuronal 15. 45

50 En la figura 1 se muestra en líneas discontinuas otra aplicación de la red neuronal entrenada 15. Una tercera red neuronal 19 entrenada con grupos 2, 12 de documentos 3, 13 en un idioma distinto en el que había sido entrenada la red neuronal 15 se asigna por nodos - si pueden identificarse nodos de clasificación correspondientes 15', 19' dentro de las redes 15 y 19 - a la segunda red 15. En las entradas y salidas 38, 39 de la tercera red neuronal 19 se utiliza otro diccionario de patrones 9, generado a partir de un grupo de documentos 2 en el idioma de la tercera red 19. De esta manera, pueden obtenerse traducciones semánticas entre dos idiomas por asignación semántica de dos redes MPF o HTM entrenadas 15, 19.

55 Aunque la invención se ha descrito con referencia a mapas 5 y patrones 6 bidimensionales, hay que señalar que la primera red neuronal 4 también podría generar mapas de tres o más dimensiones 5, lo que da lugar a patrones de tres o más dimensiones 6 en el diccionario de patrones 9, posteriormente a secuencias patrones de tres o más dimensiones 14 y una segunda y una tercera red neuronal 15, 19 que funcionan en tres o más dimensiones.

60 La invención no está limitada de ninguna manera a las realizaciones específicas descritas como ejemplos en detalle, sino que comprende todas las variantes, modificaciones y combinaciones de las mismas que están comprendidas en el alcance de las reivindicaciones adjuntas.

**REIVINDICACIONES**

1. Procedimiento implementado por ordenador para generar un diccionario legible por ordenador para traducir texto en una forma legible por una red neuronal, que comprende:

5 entrenar una primera red neuronal (4) de tipo de mapa de auto-organización con un primer grupo (2) de primeros documentos (3) cada uno conteniendo una o más palabras clave (7) en un contexto semántico, estando entrenada la primera red neuronal (4) con vectores de entrada (21) que representan cada uno un documento (3) del primer grupo (2) y su contenido de palabras clave, para  
10 asignar cada documento de texto (3) a un punto ( $X_i/Y_j$ ) en el mapa de auto-organización (5) por agrupación semántica, como resultado de cuyo entrenamiento, en el mapa (5), los documentos (3) han sido asignados a puntos individuales ( $X_i/Y_j$ ) del mapa (5);  
15 determinar, para cada palabra clave (7) que tiene lugar en el primer grupo (2), todos los puntos ( $X_i/Y_j$ ) en el mapa de auto-organización (5) al cual se asignan documentos de texto (3) que contienen dicha palabra clave (7), como un patrón de dos o más dimensiones (6) de puntos ( $X_i/Y_j$ ) asociados a dicha palabra clave (7); y  
20 guardar todas las palabras clave (7) y patrones asociados (6) como un diccionario de patrones legible por ordenador (9), estando asociado cada patrón (6) a una palabra clave (7) en el diccionario de patrones (9).

2. Procedimiento de acuerdo con la reivindicación 1 para entrenar una red neuronal, que comprende, además:

25 formar por lo menos una secuencia (11) de palabras clave (7) a partir de un segundo grupo (12) de segundos documentos de texto (13) cada uno conteniendo una o más palabras clave (7) en un contexto semántico;  
traducir dicha por lo menos una secuencia (11) de palabras clave (7) en por lo menos una secuencia (14) de patrones (6) utilizando dicho diccionario de patrones (9); y  
30 entrenar una segunda red neuronal (15) con dicha por lo menos una secuencia (14) de patrones (6).

3. Procedimiento de acuerdo con la reivindicación 2, caracterizado por el hecho de que la segunda red neuronal (15) es jerárquica y por lo menos parcialmente recurrente.

4. Procedimiento de acuerdo con la reivindicación 2, caracterizado por el hecho de que la segunda red neuronal (15) es un marco de predicción de memoria.

5. Procedimiento de acuerdo con la reivindicación 2, caracterizado por el hecho de que la segunda red neuronal (15) es una memoria temporal jerárquica.

6. Procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 5, caracterizado por el hecho de que la primera red neuronal (4) es un mapa de auto-organización de Kohonen.

7. Procedimiento de acuerdo con cualquiera de las reivindicaciones 2 a 6, caracterizado por el hecho de que, para cada uno de los segundos documentos (13) del segundo grupo (12), se forma una secuencia independiente (11) de palabras clave (7) y se traduce en una secuencia independiente (14) de patrones (6) y la segunda red neuronal (15) es entrenada sucesivamente con cada una de dichas secuencias independientes (11) de patrones (6).

8. Procedimiento de acuerdo con la reivindicación 7, caracterizado por el hecho de que los segundos documentos (13) están ordenados y, cuando se entrena la segunda red neuronal (15), las secuencias independientes (14) de patrones (6) se envían a la segunda red neuronal (15) en el orden en que están clasificados los segundos documentos (13) a partir del cual han sido formados y traducidos cada uno.

9. Procedimiento de acuerdo con la reivindicación 8, caracterizado por el hecho de que los segundos documentos están ordenados por complejidad ascendente, en el que la complejidad de un segundo documento (13) se determina en base a uno o más de: el número de palabras clave diferentes (7) en ese segundo documento (13), la duración media de una frase en ese segundo documento (13), la frecuencia de una o más palabras clave (7) del primer grupo en ese segundo documento (13), la frecuencia de una o más palabras clave (7) de ese segundo documento (13) en el primer grupo (2) u otro corpus de texto.

10. Procedimiento de acuerdo con cualquiera de las reivindicaciones 2 a 9, para el procesamiento de texto que contiene por lo menos una palabra clave, que comprende:

traducir dicha por lo menos una palabra clave (7) en por lo menos un patrón (6) por medio del diccionario de patrones (9);

enviar dicho por lo menos un patrón (6) como patrón de entrada a dicha segunda red neuronal entrenada (15);  
obtener por lo menos un patrón de salida (6) a partir de dicha segunda red neuronal entrenada; y  
traducir dicho por lo menos un patrón de salida (6) en por lo menos una palabra clave (7) por medio del diccionario de patrones (9).

5  
10 11. Procedimiento de acuerdo con la reivindicación 10, para la clasificación semántica de texto, caracterizado por el hecho de que la segunda red neuronal (15) es jerárquica, dicho por lo menos un patrón de entrada (6) se envía a por lo menos una capa inferior de la jerarquía y dicho por lo menos un patrón de salida (6) se obtiene a partir de por lo menos una capa superior de la jerarquía.

15 12. Procedimiento de acuerdo con la reivindicación 10, para la predicción semántica de texto, caracterizado por el hecho de que la segunda red neuronal (15) es jerárquica, dicho por lo menos un patrón de entrada (6) se envía a por lo menos una capa superior de la jerarquía y dicho por lo menos un patrón de salida (6) se obtiene a partir de por lo menos una capa inferior de la jerarquía.

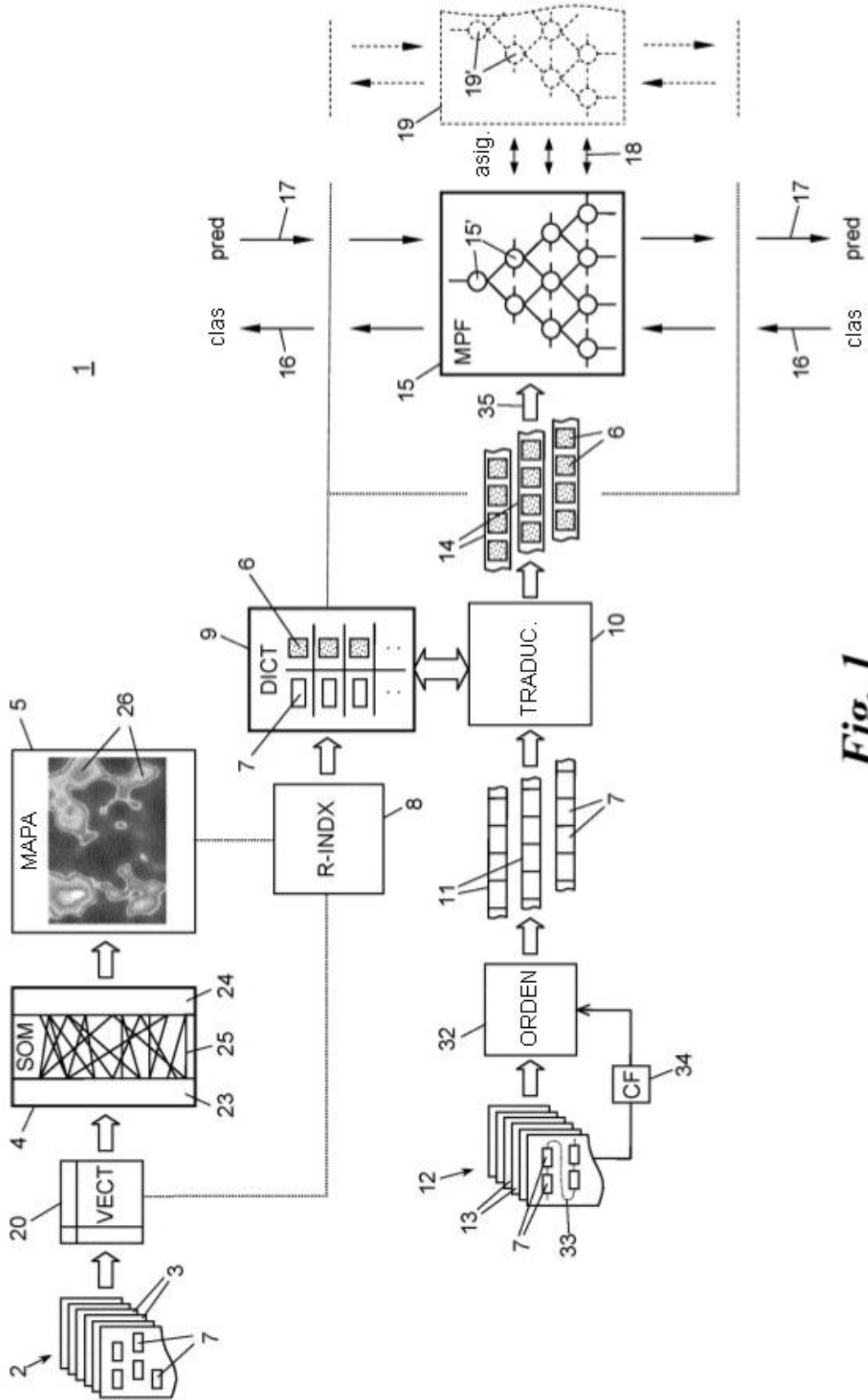
13. Diccionario legible por ordenador realizado en un medio legible por ordenador, generado con un procedimiento de acuerdo con la reivindicación 1.

20 14. Máquina de clasificación o predicción, que comprende una red neuronal de tipo jerárquico que ha sido entrenada como dicha segunda red neuronal (15) con un procedimiento de acuerdo con una de las reivindicaciones 2 a 9.

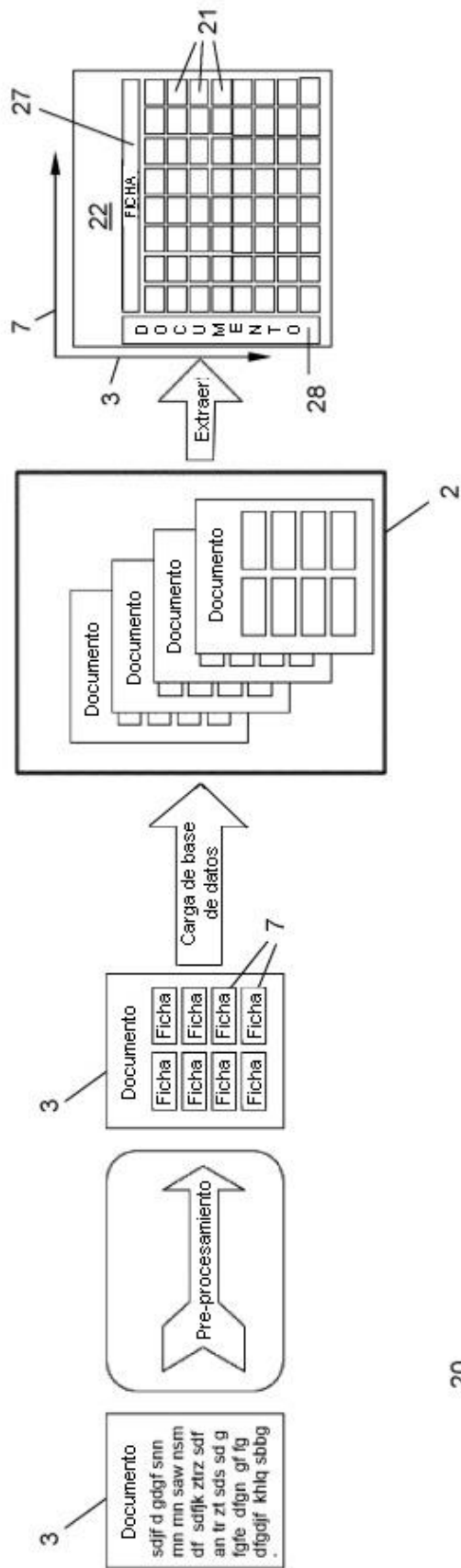
25 15. Máquina de traducción, que comprende una máquina de clasificación de acuerdo con la reivindicación 14, cuya red neuronal (15) ha sido entrenada con un procedimiento de acuerdo con una de las reivindicaciones 2 a 9 utilizando primeros y segundos documentos de texto (3, 13) en un primer idioma; y

una máquina de predicción de acuerdo con la reivindicación 14, cuya red neuronal (19) ha sido entrenada con un procedimiento de acuerdo con una de las reivindicaciones 2 a 9 utilizando primeros y segundos documentos de texto (3, 13) en un segundo idioma;

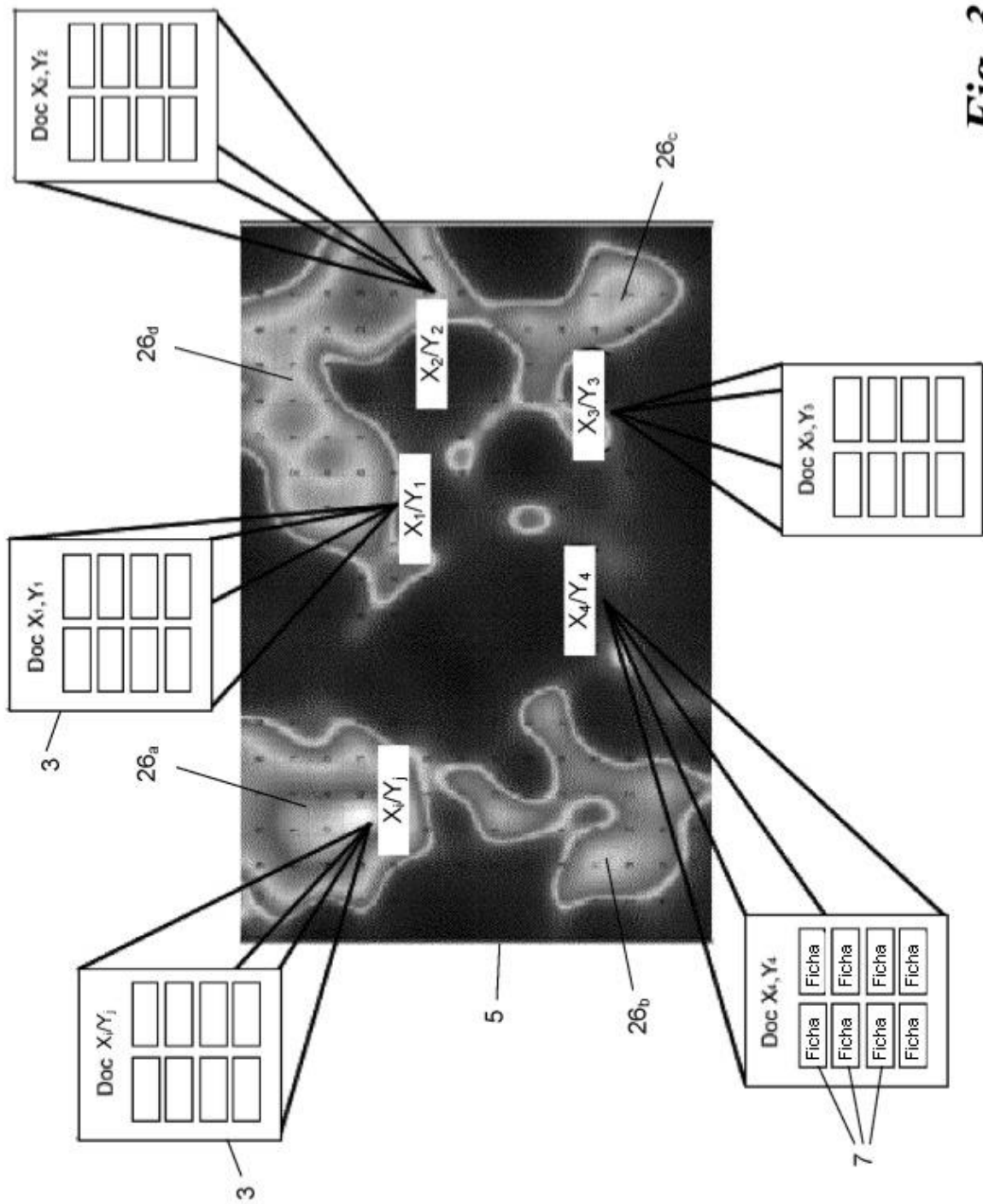
30 en el que nodos (15') de la red neuronal (15) de la máquina de clasificación están conectados a nodos (19') de la red neuronal (19) de la máquina de predicción.



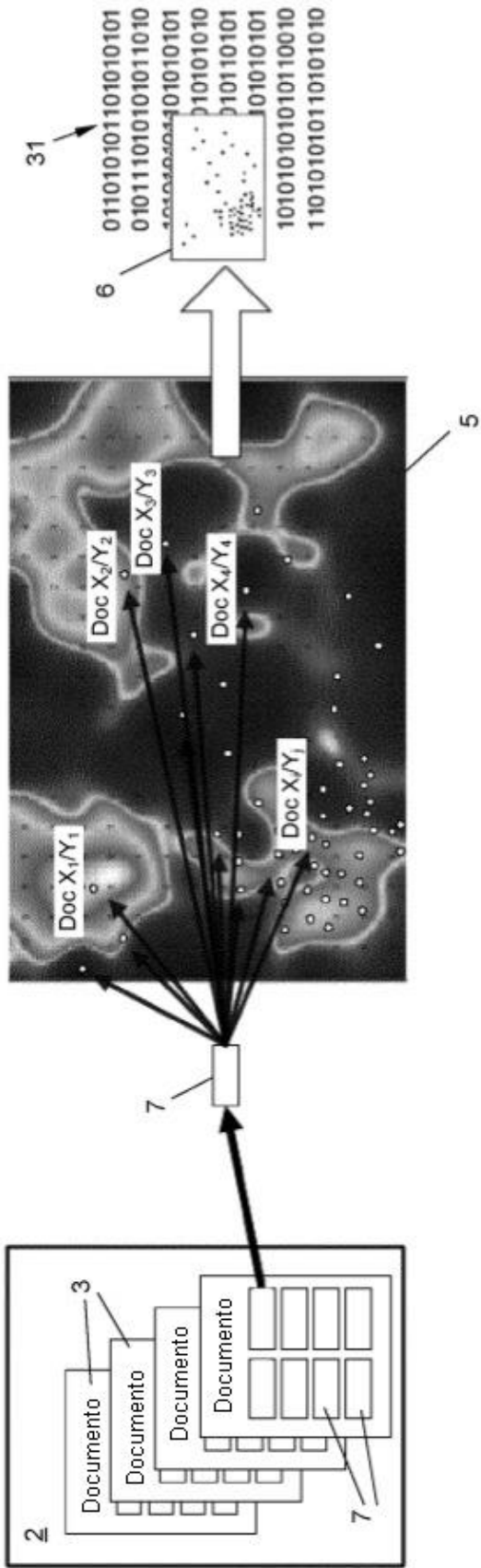
**Fig. 1**



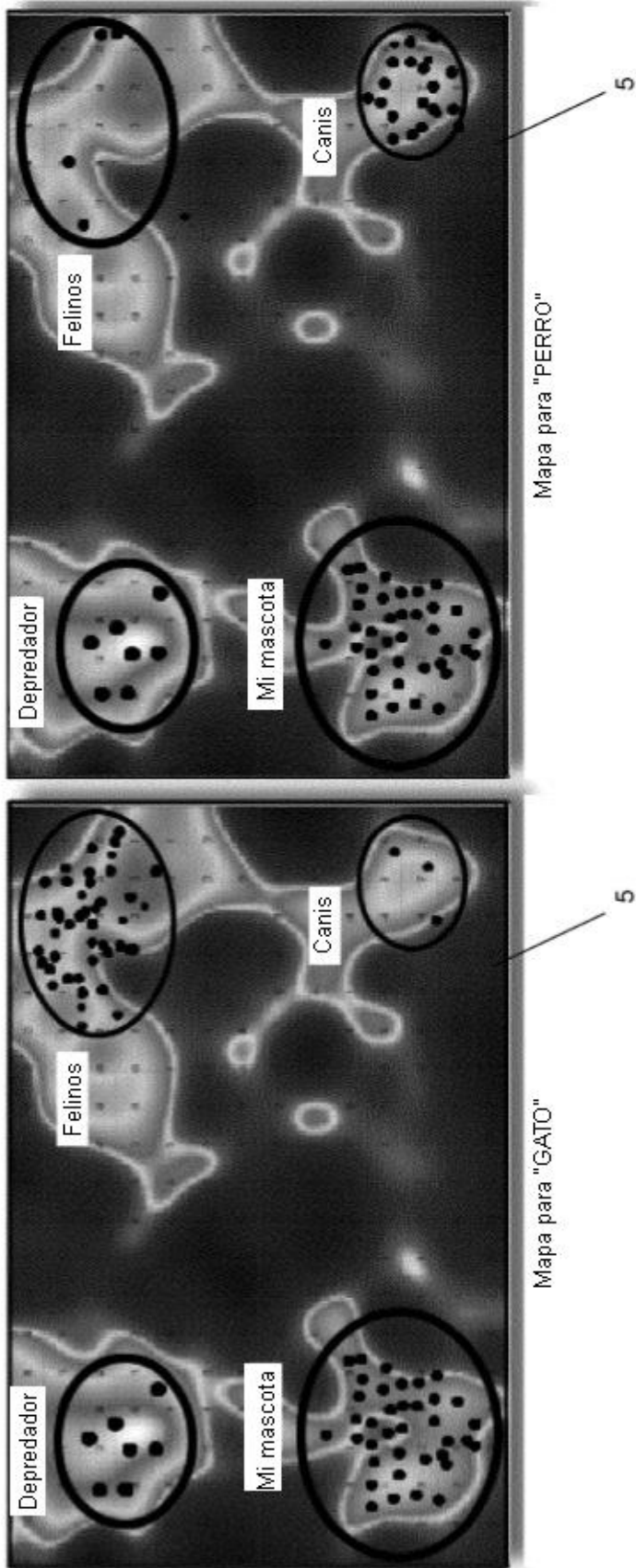
**Fig. 2**



**Fig. 3**

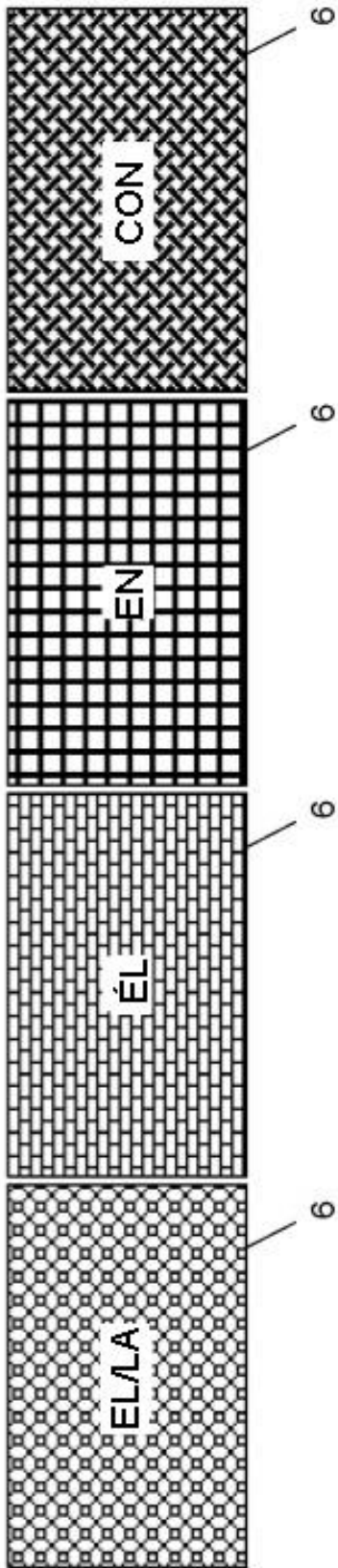


**Fig. 4**



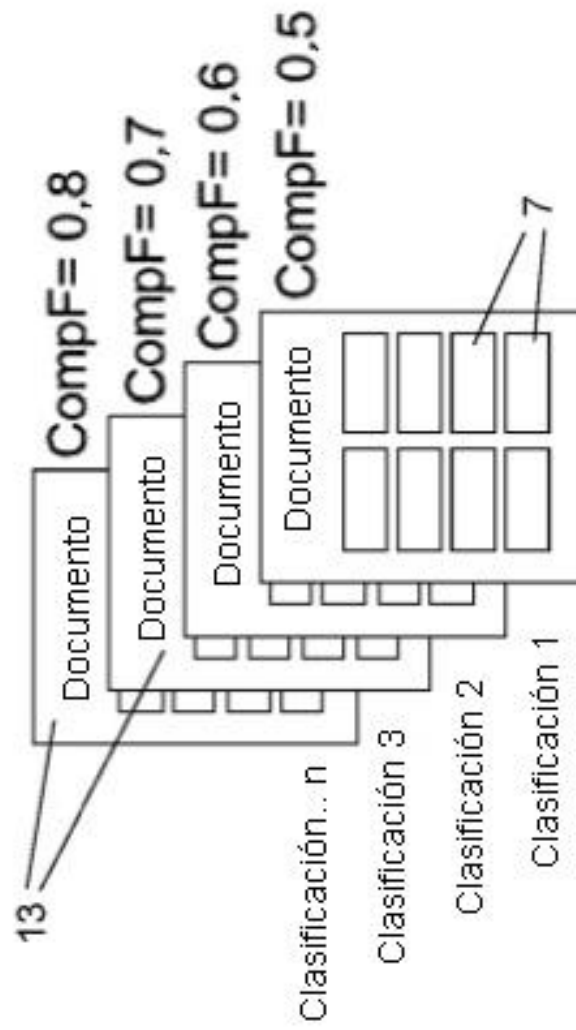
**Fig. 5**





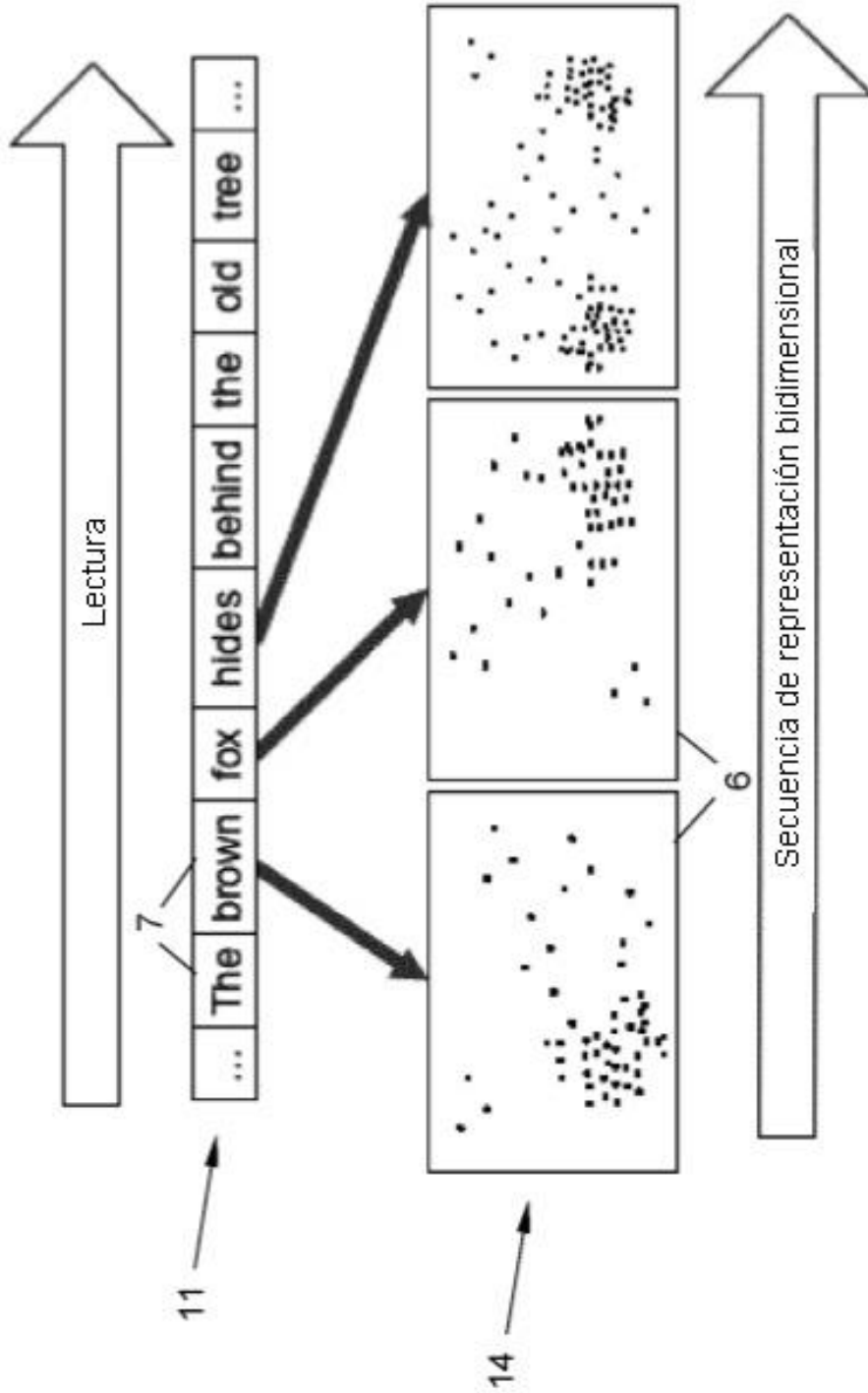
*Fig. 6*





**Fig. 8**

10



**Fig. 9**

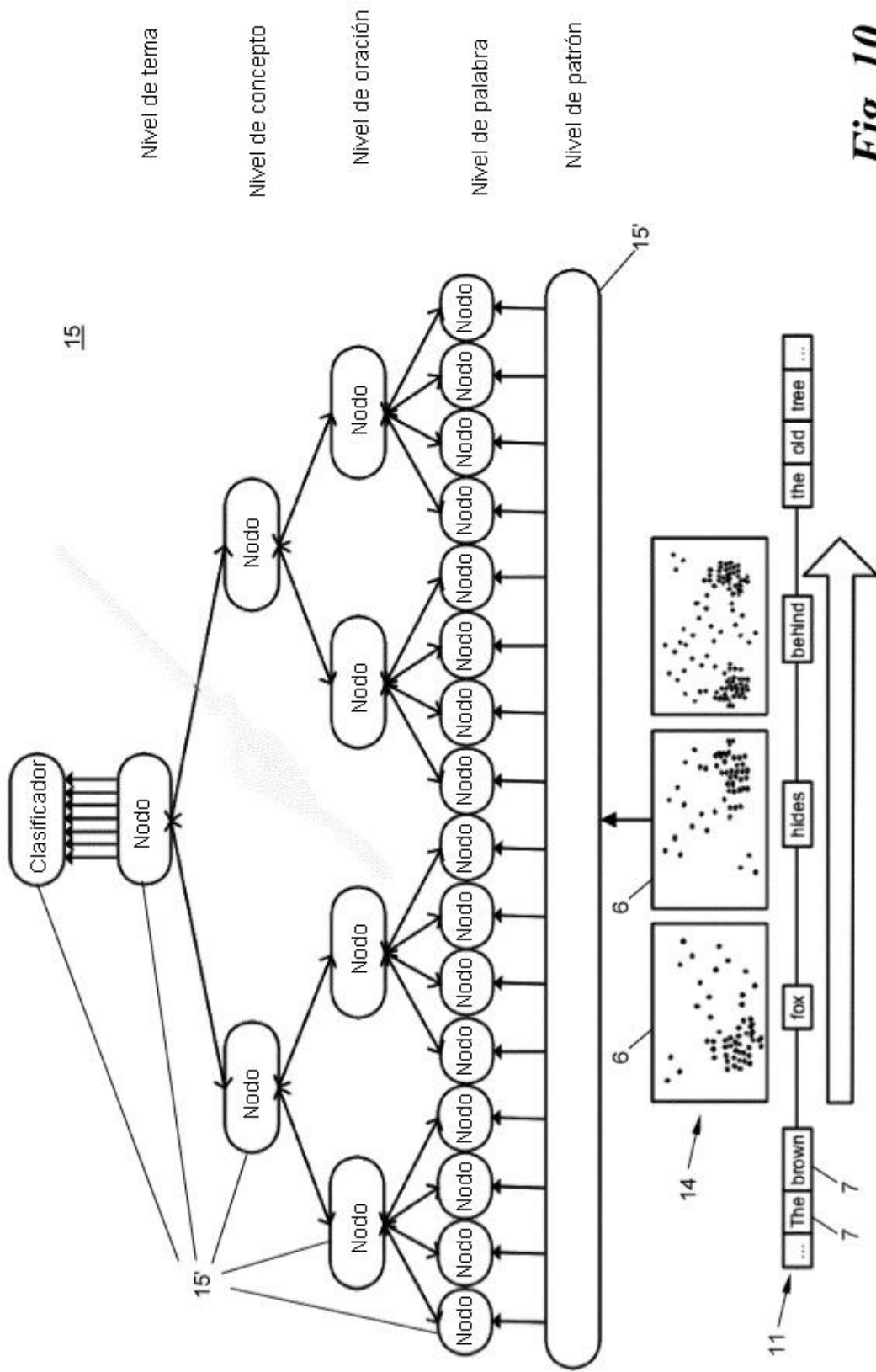


Fig. 10

## REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante es únicamente para la comodidad del lector. No forma parte del documento de la patente europea. A pesar del cuidado tenido en la recopilación de las referencias, no se pueden excluir errores u omisiones y la EPO niega toda responsabilidad en este sentido.

## Documentos de patentes citados en la descripción

- 10 • US 20070276774 A1 [0020] • US 20110225108 A1 [0020]  
 • US 20080059389 A1 [0020] • US 8037010 B2 [0020]  
 • US 7739208 B2 [0020] • US 8103603 B2 [0020]  
 • US 7937342 B2 [0020]

## Literatura diferente de patentes citada en la descripción

- 15 • **BISHOP C. M.** Neural Networks for Pattern Recognition. Oxford University Press, 1995 [0002] 40  
 • **REY, G. D. ; WENDER K. F.** Neurale Netze. Hans Huber, 2011 [0002]  
 20 • A Latent Semantic Indexing Approach to Self-Organizing Maps of Document Collections. **AMPAZIS N. et al.** Neural Processing Letters. Kluwer Academic Publishers, 2004, 1-17 [0008] 45  
 • **KOHONEN, T.** The Self-Organizing Map. Proceedings of the IEEE, 1990, vol. 78 (9), 1464-1480 [0018] 45  
 • **KOHONEN, T. ; SOMERVUO, P.** Self-Organizing Maps of Symbol Strings. Neurocomputing, 1998, vol. 21 (1-3), 19-30 [0018] 50  
 • **KASKI, S. ; HONKELA, T. ; LAGUS, K. ; KOHONEN, T.** Websom-Self-Organizing Maps of Document Collections. Neurocomputing, 1998, vol. 21 (1-3), 101-117 [0018] 55  
 • **MERKL, D.** Text Classification with Self-Organizing Maps: Some Lessons Learned. Neurocomputing, 1998, vol. 21 (1-3), 61-77 [0018] 55  
 35 • **VESANTO, J. ; ALHONIEMI, E.** Clustering of the Self-Organizing Map. *IEEE Transactions on Neural Networks*, 2000, vol. 11 (3), 586-600 [0018] 60  
 • **PÖLZLBAUER G. ; DITTENBACH M. ; RAUBER A.** Advanced Visualization of Self-Organizing Maps with Vector Fields *IEEE Transactions on Neural Networks* 2006, vol. 19, 911-922 [0018] 65
- **HAWKINS, J. ; GEORGE, D. ; NIEMASIK, J.** Sequence Memory for Prediction, Inference and Behaviour. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 2009, vol. 364 (1521), 1203-9 [0020]  
 • **STARZYK, J. A. ; HE, H.** Spatio-Temporal Memories for Machine Learning: A Long-Term Memory Organization. *IEEE Transactions on Neural Networks*, 2009, vol. 20 (5), 768-80 [0020]  
 • Hierarchical Temporal Memory Including HTM Cortical Learning Algorithms. Whitepaper of Numenta, Inc, 12 September 2011 [0020]  
 • **RODRIGUEZ A. ; WHITSON J. ; GRANGER R.** Derivation and Analysis of Basic Computational Operations of Thalamocortical Circuits. *Journal of Cognitive Neuroscience*, 2004, vol. 16 (5), 856-877 [0020]  
 • **RODRIGUEZ, R. J. ; CANNADY, J. A.** Towards a Hierarchical Temporal Memory Based Self-Managed Dynamic Trust Replication Mechanism in Cognitive Mobile Ad-hoc Networks. *Proceedings of the 10th WSEAS international conference on artificial intelligence, knowledge engineering and data bases*, 2011 [0020]