

Aprendizaje Automático aplicado a Reconocimiento de Patrones en Video y Minería de Datos

Laura Lanzarini¹, Waldo Hasperué¹, César Estrebou¹, Franco Ronchetti^{1,2}, Augusto Villa Monte^{1,2}, Germán Aquino^{1,3}, Facundo Quiroga^{1,2}, María José Basgall^{1,3}, Luis Rojas⁴, Patricia Jimbo Santana⁵
Aurelio Fernandez Bariviera⁶, José Ángel Olivas Varela⁷

¹ Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática - Universidad Nacional de La Plata

² Becario postgrado UNLP ³ Becario postgrado CONICET

⁴ Facultad de Ingeniería. Universidad Nacional de la Patagonia San Juan Bosco

⁵ Dpto. Ciencias de la Computación, Universidad de las Fuerzas Armadas ESPE, Ecuador

⁶ Facultat de Ciències Econòmiques i Empresariales, Universitat Rovira I Virgili Tarragona, España

⁷ Escuela Superior de Informática, Universidad de Castilla-La Mancha, España

{laural, whasperue, cesarest, fronchetti, avillamonte, gaquino, fquiroga, mjbasgall}@lidi.info.unlp.edu.ar
luisf.09@gmail.com, pjimbo@pepsolutions.com, aurelio.fernandez@urv.net, joseangel.olivas@uclm.es

CONTEXTO

Esta presentación corresponde al Subproyecto “Sistemas Inteligentes” perteneciente al proyecto “Cómputo paralelo de altas prestaciones. Fundamentos y evaluación de rendimiento en HPC. Aplicaciones a sistemas inteligentes, simulación y tratamiento de imágenes” (Periodo 2014–2017) del Instituto de Investigación en Informática LIDI.

RESUMEN

Esta línea de investigación se centra en el estudio y desarrollo de Sistemas Inteligentes basados en Redes Neuronales Artificiales aplicadas al Reconocimiento de Patrones y a la Minería de Datos.

Con respecto al Reconocimiento de Patrones, el énfasis está puesto en el análisis de videos con el objetivo de identificar acciones humanas que faciliten la interfaz hombre/máquina y en la detección de patrones de movimiento en videos provenientes de cámaras de vigilancia asociados con situaciones de riesgo.

En el área de la Minería de Datos se está trabajando en la generación de un modelo de fácil interpretación a partir de la extracción de reglas de clasificación que permita justificar la toma de decisiones y en desarrollo de nuevas estrategias aplicables a flujos de datos.

Con respecto a Minería de Textos se han desarrollado métodos capaces de extraer las palabras clave de documentos independientemente del lenguaje. Actualmente se están desarrollando estrategias para resumir documentos a través de la extracción de los párrafos más representativos utilizando técnicas de optimización.

Palabras clave: Redes Neuronales, Técnicas de Optimización, Estrategias adaptativas, Reconocimiento de Patrones, Minería de Datos, Minería de Textos.

1. INTRODUCCION

El Instituto de Investigación en Informática LIDI tiene una larga trayectoria en el estudio, investigación y desarrollo de Sistemas Inteligentes basados en distintos tipos de estrategias adaptativas. Los resultados obtenidos han sido medidos en la solución de problemas pertenecientes a distintas áreas. A continuación se detallan los resultados obtenidos durante el último año.

1.1. Reconocimiento de Patrones

En el III LIDI, desde hace varios años se vienen trabajando en procesamiento de señales de audio y video. Como resultado de estas investigaciones se han diseñado e

implementado técnicas originales aplicables al reconocimiento de distintas medidas biométricas así como de gestos dinámicos. Continuando con esta línea, actualmente se están desarrollando los siguientes temas:

Reconocimiento de gestos

La aparición de nuevas tecnologías en sensores y la popularidad de los dispositivos móviles han introducido nuevas posibilidades de interacción hombre-máquina y obtención de información. No obstante, la habilidad de utilizar efectivamente estos datos y modos de comunicación requiere novedosas técnicas de procesamiento, análisis, e inferencia a partir de señales.

En esta línea de investigación, se ha trabajado en la búsqueda de descriptores para acciones humanas dinámicas que permitan, junto con un modelo de clasificación adecuado, un reconocimiento efectivo de las acciones capturadas en videos [1].

Luego de explorar el reconocimiento de gestos y de acciones humanas en general [1,2], en el III-LIDI se ha re-enfocado el objetivo de la investigación en técnicas de reconocimiento de señales temporales complejas para obtener reconocedores de lenguaje de señas robustos.

Se crearon dos bases de datos: la primera, de configuraciones de manos del lenguaje de señas argentino y la segunda de señas dinámicas que utilizan esas configuraciones. Su creación tiene el objetivo de poder realizar experimentos con distintos métodos de reconocimiento y de posibilitar la creación de software que traduzca automáticamente desde el video al lenguaje castellano.

En esta línea de investigación, se publicaron los detalles de la base de datos de configuraciones, junto con un experimento de desempeño inicial que compara métodos tradicionales del estado del arte con un modelo que utiliza descriptores basados en la transformada de Radón y el ProbSom como clasificador [3]. Actualmente, se está trabajando en la clasificación de las señas dinámicas de la segunda base de datos.

Detección de patrones en movimiento en video

Esta línea se ha centrado principalmente en el procesamiento y reconocimiento de patrones en señales de audio e imágenes con el fin de identificar a una persona a través de su rostro o el sonido de su voz. Los resultados obtenidos se encuentran publicados en [4,5,6].

En el último tiempo, se ha comenzado a estudiar la posibilidad de incorporar estos resultados al análisis de videos. Se espera poder contribuir al diseño y desarrollo de nuevas estrategias adaptativas aplicables al análisis la escena. Interesa especialmente investigar aquellas estrategias que permitan reconocer automáticamente la estructura de la escena, las actividades y los patrones de movimiento con el objetivo de cambiar el seguimiento de objetos de bajo nivel por el análisis de eventos de alto nivel y la detección de situaciones anómalas.

Los resultados obtenidos serán utilizados en la detección automática de situaciones de riesgo o amenazas en escenas captadas a través de sistemas de video-vigilancia.

1.2. MINERÍA DE DATOS

Obtención de Reglas de Clasificación

Esta línea de investigación está centrada en la obtención de reglas de clasificación, del tipo IF-THEN, a partir de redes neuronales y técnicas de optimización.

En especial se estudian métodos de clustering y clasificación de patrones que permitan identificar aquellos atributos relevantes para el problema. Dichos atributos serán especialmente considerados en el momento de construir el antecedente de la regla. Reducir la cantidad de atributos a considerar ayuda a obtener un conjunto de reglas sencillas de baja cardinalidad, facilitando de esta forma el criterio que da lugar a la clasificación.

Como área de transferencia tecnológica se ha analizado la situación de las compañías financieras al momento de determinar el riesgo en una operación de otorgamiento de crédito. En estos casos, el proceso de toma de

decisiones habitual consiste en la recopilación de información personal y financiera sobre el prestatario con el objetivo de cuantificar el riesgo que corre la empresa otorgándole el dinero solicitado. Este enfoque no siempre ofrece buenos resultados.

La aplicación del método definido en [7] fue mejorado en [8] al reemplazar la red neuronal SOM por una red neuronal LVQ que por ser supervisada da lugar a una mejor inicialización de la técnica de optimización que determina el conjunto de reglas.

Luego en [9] se utilizó esta nueva estrategia para obtener reglas asociadas a riesgo crediticio demostrando que posee una reducción en el error de Tipo I mejorando de esta forma la identificación de los futuros clientes de la compañía.

Procesamiento de Flujo de Datos

Esta línea de investigación está centrada en el procesamiento de grandes volúmenes de tweets. Los temas abordados en esta línea comprenden el procesamiento del lenguaje natural, la detección de tópicos y el análisis de sentimiento.

Dado el gran volumen de tweets que se genera constantemente se están desarrollando estrategias de procesamiento que involucran el análisis de flujo continuo de tweets otorgando una respuesta rápida la cual se adapta dinámicamente ante la llegada de nuevos tweets.

Se están desarrollando técnicas de procesamiento distribuido en el paradigma MapReduce. Estas técnicas están enfocadas en la tarea de clustering dinámico. El objetivo de las mismas es poder variar los centros dinámicamente agregando y eliminando elementos (palabras) a medida que se detecta la aparición de nuevos tópicos, establecer relaciones entre los tópicos en un momento dado como así también determinar el instante en que determinadas palabras dejan de utilizarse [10].

1.3. MINERIA DE TEXTOS

Extracción de palabras clave en documentos de texto

Esta línea de investigación tiene su eje central en el estudio y aplicación de distintos métodos de representación de documentos así como de distintas técnicas adaptativas aplicables en la resolución de problemas de extracción de palabras clave, tarea de sumo interés ya que permite caracterizar un documento facilitando su búsqueda y clasificación.

En esta línea, se desarrolló un método de identificación de palabras clave a partir de documentos de texto en español utilizando redes neuronales inicialmente descrito en [11] y posteriormente mejorado a través del uso de redes auto-asociativas o autoencoders, los cuales son adecuados en problemas de clasificación muy desbalanceados.

Este método define una representación vectorial para los términos, para luego aplicar un proceso de filtrado gramatical con el fin de remover términos inválidos. Es importante destacar que ésta es la única parte del método que es dependiente del idioma en cuestión.

Una vez obtenida la representación se utiliza un ensemble de autoencoders combinados con redes neuronales convencionales para construir un modelo de clasificación. Dado un término de un documento nuevo los autoencoders determinan para el mismo un error de reconstrucción, que es utilizado para definir un ranking de términos.

Con el objeto de medir la eficacia del método propuesto se utilizaron artículos científicos de WICC y CACIC. Los resultados obtenidos publicados en [12] permiten afirmar que el método propuesto es comparable o incluso superior a Maui, uno de los algoritmos más reconocidos en esta problemática.

Actualmente se continúa el desarrollo del método para mejorar su precisión y ampliar su dominio de aplicación.

Síntesis automática de documentos

En este trabajo de investigación se busca identificar y representar los temas relevantes de un documento. Esto resulta de sumo interés para que a partir de uno o varios textos se pueda comprender la temática general tratada en el mismo, obtener las ideas centrales, categorizarlos automáticamente o comparar su contenido. Anteriormente se desarrolló una estrategia para medir la relevancia de los términos más utilizados en un conjunto de e-mails [13] que permitió mejorar la atención de los alumnos.

Actualmente se está trabajando en la representación de una determinada colección de documentos de texto con el objetivo de identificar las partes relevantes de cada uno de ellos. Para ello se eligió una serie de tesis escritas en LaTeX con temas en común. Se las procesó de manera de tener cada sentencia identificada por su contenido y ubicación dentro del documento, permitiendo que se les calcule un conocido conjunto de métricas. Por otro lado para cada sentencia se estableció la puntuación considerada al momento de construir la síntesis del documento.

A partir de la representación obtenida, se están evaluando distintas técnicas de optimización basadas en cúmulos de partículas con el fin de identificar el aporte de cada métrica en la puntuación propuesta. Esto permite encontrar la combinación adecuada de métricas que dan lugar al resumen esperado de cada documento.

2. TEMAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio de técnicas de optimización poblaciones y redes neuronales artificiales para la obtención de reglas de tipo IF-THEN.
- Clasificación de configuraciones de manos para el lenguaje de señas.
- Desarrollo de representaciones para el reconocimiento de señas.
- Métodos estructurados y no estructurados

a la representación de documentos.

- Problemas de clasificación con desbalance de clase severo.
- Síntesis de documentos de texto a partir de la selección de las partes relevantes obteniendo la combinación de un conjunto de métricas.
- Implementación de técnicas para el tratamiento de flujos continuos de tweets para la obtención de respuestas dinámicas.
- Implementación de técnicas en el paradigma de MapReduce para clustering dinámico.

3. RESULTADOS OBTENIDOS

- Desarrollo de un método de obtención de reglas de clasificación con énfasis en la reducción de la complejidad del modelo aplicable a riesgo crediticio.
- Elaboración de una base de datos de configuraciones de manos del Lenguaje de Señas Argentino, y evaluación de su desempeño con varios métodos.
- Elaboración de una base de datos de señas dinámicas del Lenguaje de Señas Argentino.
- Desarrollo de una representación de términos que junto con un modelo de clasificación permite identificar palabras clave en un documento.
- Implementación de una técnica para la tarea de clustering dinámico en el paradigma MapReduce.
- Análisis de sentimiento de tweets relacionados con los tópicos del cambio climático en relación a la cumbre de París (COP21) realizada en noviembre-diciembre de 2015. Esta tarea se está realizando conjuntamente con integrantes del departamento de Educación de la Universidad de Stanford.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de los temas involucrados en esta línea de investigación se han finalizado 2 tesis de doctorado, 2 tesis de maestría y 3 tesis de especialista y numerosas tesinas de grado de Licenciatura.

Actualmente se están desarrollando 6 tesis de doctorado, 1 de maestría y 2 tesinas de grado de Licenciatura. También participan en el desarrollo de las tareas becarios y pasantes del III-LIDI.

5. REFERENCIAS

- [1] Ronchetti F., Quiroga F., Lanzarini L., Estrebou C. *Distribution of action movements (DAM): a descriptor for human action recognition*. Frontiers of Computer Science. Springer - Higher Education Press. Pp 1-10. ISSN 2095-2228. Septiembre 2015
- [2] Quiroga F., Corbalán L. *A novel competitive neural classifier for gesture recognition with small training sets*. XVIII Congreso Argentino de Ciencias de la Computación. 2013. pp. 140-149.
- [3] Ronchetti F., Quiroga F., Estrebou C., Lanzarini L. *Clasificación de configuraciones de manos del Lenguaje de Señas Argentino con ProbSOM*. XXI Congreso Argentino de Ciencias de la Computación. CACIC 2015.
- [4] Maulini J., Lanzarini L. *Face Recognition using SIFT descriptors and Binary PSO with velocity control*. Computer Science & Technology Series – XVII Argentine Congress of Computer Science Selected Papers. ISBN 978-950-34-0885-8. EDULP, 2012. pp. 43-53.
- [5] Albanesi B., Funes N., Chichizola F., Lanzarini L. *Reconocimiento de Objetos en Video utilizando SIFT paralelo*. XVI Congreso Argentino de Ciencias de la Computación, CACIC 2010.
- [6] Estrebou C., Lanzarini L., Hasperué W. *Voice recognition based on probabilistic SOM*. Conferencia Latinoamericana de Informática. CLEI 2010.
- [7] Lanzarini, L., Villa Monte, A., Ronchetti, F.: *SOM+PSO. A Novel Method to Obtain Classification Rules*. Journal of Computer Science & Technology (JCS&T), Vol. 15, No 1, pp. 15-22. ISSN 1666-6038. Abril 2015.
- [8] Lanzarini L., Villa Monte A., Aquino G., De Giusti A. *Obtaining classification rules using lvqPSO*. Advances in Swarm and Computational Intelligence. Lecture Notes in Computer Science. Vol 6433, pp. 183-193. ISSN 0302-9743. Springer-Verlag Berlin Heidelberg. Junio 2015.
- [9] Lanzarini L., Villa Monte A., Fernandez Bariviera A., Jimbo Santana P. *Obtaining Classification Rules Using LVQ+PSO: an application to Credit Risk*. Scientific Methods for the Treatment of Uncertainty in Social Sciences. Advances in Intelligent Systems and Computing. Springer-Verlag Berlin Heidelberg. vol. 377. Pp 383-391. ISSN 2194-5357. 2015.
- [10] Basgall M. J., Aquino G., Lanzarini L., Naiouf M. "Un Enfoque Dinámico para la Detección de Relaciones entre Tópicos en Textos provenientes de Redes Sociales". III Jornadas de Cloud Computing y Big Data (JCC&BD). Facultad de Informática. UNLP. 29 de junio al 3 de Julio de 2015.
- [11] Aquino, G., Hasperué, W., Estrebou, C., Lanzarini, L. *A Novel Language-Independent Keyword Extraction Method*. Publicado en el Libro *Computer Science & Technology Series – XIX Argentine Congress of Computer Science - Selected Papers*. Editores: Jorge Raúl Finocchietto, Patricia Pesado. EDULP, 2014. Pags.221-232.
- [12] Aquino G., Lanzarini L. *Keyword Identification in Spanish Documents using Neural Networks*. XXI Congreso Argentino de Ciencias de la Computación. CACIC 2015.
- [13] Villa Monte, A. Estrebou, C., Lanzarini, L. *E-mail processing using data mining techniques*. Computer Science & Technology Series – XVI Argentine Congress of Computer Science Selected Papers, ISBN 978-950-34-0757-8. EDULP, 2011. pp. 109-120.