

TOMA DE DECISIONES EN SISTEMAS DE EVENTOS MEDIANTE INTELIGENCIA COMPUTACIONAL Y COMPUTACIÓN DE ALTO DESEMPEÑO

Esteban Schab^(1,3), Mariela Lopresti⁽²⁾, Natalia Miranda⁽²⁾,
Carlos Casanova^(1,3) y Fabiana Piccoli^(1 y 2)

⁽¹⁾ Universidad Autónoma de Entre Ríos, Concepción del Uruguay

⁽²⁾ LIDIC- Univ. Nacional de San Luis, San Luis

⁽³⁾ Univ. Tecnológica Nacional, Facultad Regional Concepción del Uruguay, Concepción del Uruguay
Argentina

{omlopres,nemiran, mpiccoli}@unsl.edu.ar
{schabe, casanovac}@frcu.utn.edu.ar

RESUMEN

La toma de decisiones en contextos dominados por grandes volúmenes de datos debe conjugar dos atributos usualmente contrapuestos: calidad y velocidad. La disponibilidad de información generada por personas y dispositivos abre nuevos desafíos en el diseño de mecanismos para aprovecharla. Estos deben ser capaces de determinar las decisiones de mayor utilidad sujetas a ventanas temporales que garanticen su factibilidad. Uno de estos mecanismos lo constituyen las distintas analíticas, las cuales buscan transformar los datos en información a través de técnicas diversas. En este trabajo proponemos dos líneas de investigación, una enfocada en la analítica prescriptiva, capaz de determinar acciones a ser ejecutadas en el momento (decisiones operativas) o en el futuro (decisiones tácticas para corto y mediano plazo, decisiones estratégicas para largo plazo) para lograr un objetivo deseado; la otra relacionada con las técnicas de aprendizaje supervisado y la recuperación de información no estructurada. En ambos, se propone la utilización de desarrollos provenientes de la Inteligencia Computacional y de la Computación de Alto Desempeño con el fin de obtener, de forma colaborativa, calidad y velocidad en las decisiones.

Palabras clave: Inteligencia Computacional. Analíticas. Big Data. Computación de Alto Desempeño.

CONTEXTO

Esta propuesta de trabajo se lleva a cabo dentro de los proyectos de investigación: “Tecnologías Avanzadas aplicadas al Procesamiento de Datos Masivos” (LIDIC, UNSL) y “Cómputo de Altas Prestaciones

aplicado a la Solución de Grandes Problemas” (UADER).

1. INTRODUCCIÓN

La mejora continua y adaptativa de los procesos de negocio resulta clave para mantener la competitividad de las organizaciones. En este contexto, la digitalización de los procesos, así como el incremento en las tecnologías de monitoreo, han llevado a producir una enorme cantidad de datos, los cuales tienen un gran potencial para mejorar los procesos conducidos por analíticas [9][12][30].

Las analíticas buscan transformar los datos en conocimiento para la toma de decisiones [11], y se distinguen cuatro tipos de analíticas según el nivel de automatización del proceso [17]. Ellas son:

- Descriptivas: intentan responder qué ha pasado o está pasando.
- Diagnósticas: por qué ha pasado o está pasando, analizando para ello datos históricos [26].
- Predictivas: buscan responder qué sucederá, aplicando el conocimiento para predecir nuevos datos sobre el presente o el futuro (pronóstico) [6].
- Prescriptivas: responden qué debería hacerse para lograr un objetivo, esto es, determina acciones a ser ejecutadas en el momento (decisiones operativas) o en el futuro (decisiones tácticas para corto y mediano plazo, decisiones estratégicas para largo plazo).

Cabe aclarar que ninguno de los tres primeros enfoques sugiere acciones concretas, sino que descansan en el juicio subjetivo y las habilidades analíticas del usuario para deducir acciones de mejora [26] [27].

Si bien el volumen de datos generados por personas y dispositivos se encuentra en continuo crecimiento, y a pesar de los avances tecnológicos, en general, las analíticas de procesos existentes dentro de la industria actual no aprovechan completamente el conocimiento oculto debido a las siguientes limitaciones [9]:

- a. No hacen uso de técnicas prescriptivas para transformar los resultados del análisis en acciones de mejora concretas, dejando este paso completamente a criterio del usuario.
- b. Hacen un uso intensivo de datos de sistemas en producción, generando un deterioro en el desempeño de las herramientas de software que soportan los procesos.
- c. La optimización es conducida *ex post*, después de completado el proceso, en contraste a la mejora proactiva durante la ejecución del proceso.

En el área de datos masivos (o *Big Data*), se identifica como área emergente el procesamiento de *datastreams*, también llamado *Data Stream Mining* [2][14][24]. Un *datastream* es una representación digital y transmisión continua de datos, los cuales describen una clase de eventos relacionada [23] [32]. Mediante el procesamiento de estos *datastreams* se puede lograr la respuesta en tiempo real a los eventos en forma de toma de decisiones.

Los grandes volúmenes de datos generados pueden ser utilizados en algoritmos de diversa índole para la generación de analíticas, particularmente las descriptivas y predictivas, utilizando por ejemplo aprendizaje supervisado o no supervisado. Existe, sin embargo, una dificultad en lo relativo a las analíticas prescriptivas: no cuentan con un “profesor” [25] que les enseñe qué acción tomar en cada circunstancia. Un tipo de aprendizaje que no necesita de un profesor es el llamado aprendizaje por refuerzo. En este esquema es el propio agente quien es capaz de juzgar y criticar sus acciones con base en sus percepciones y de alguna medida de aptitud, recompensa o refuerzo. La tarea del aprendizaje por refuerzo es usar recompensas observadas para aprender una política óptima (o aproximadamente óptima) del entorno, sin asumir ningún conocimiento *a priori* [25][29]. Esta política le dice al agente qué hacer en cada estado posible a alcanzar.

Por otra parte, la Inteligencia Computacional (IC), que toma su inspiración de la naturaleza [13], apunta a resolver los problemas aprovechando la imprecisión y la incertidumbre presente en el proceso de toma de decisiones brindando las “soluciones más satisfactorias”. Por ello, el aprendizaje por refuerzo en combinación con técnicas de IC son un activo fundamental para la elaboración de analíticas prescriptivas que permitan resolver problemas concernientes a la toma de decisiones en contextos de incertidumbre.

Cualquiera sea la analítica, su generación debe ser lo suficientemente rápida como para procesar los *datastreams* generados en forma continua por el sistema, esto implica el uso de técnicas y herramientas de Computación de Alto Desempeño (HPC).

En la siguiente sección se describen con mayor detalle las características de las líneas de investigación.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

En este contexto de Procesos de negocio y su mejora, los *datastreams* abren nuevas y amplias oportunidades para la creación de valor en las organizaciones. A través de su procesamiento se puede conseguir respuesta en tiempo real a eventos en forma de toma de decisiones.

En [8][10][32] encontramos distintos enfoques para la definición de eventos en un proceso de negocio, variando en su complejidad. Una secuencia o flujo de eventos (*Event Stream*) es una serie ordenada y potencialmente ilimitada de eventos [21]. Los flujos de eventos se generan y utilizan en muchos sistemas. Ejemplo de ellos son:

- Bancos y otros sistemas de atención al público: En ellos la generación de *datastreams* es causada por los sistemas de gestión de atención implementados [19]. Estos sistemas en general utilizan modelos relacionales de bases de datos y, si bien permiten la elaboración de analíticas, puede resultar inapropiada su implementación. En un contexto de *Software as a Service* (SaaS) [31], el desempeño del sistema de atención, puede colapsar ante las continuas consultas

por monitoreo. Por tal motivo, la generación y el procesamiento en paralelo al sistema de atención de los eventos como *datastreams*, puede ser la tecnología de base para un monitoreo eficiente.

- Enrutamiento de vehículos: conocido por sus siglas VRP (*Vehicle Routing Problem*) [3], con suministro de información y reencaminamiento en tiempo real, orientado a la búsqueda de un paradigma de movilidad inteligente [16]. Dentro de este problema se pueden estudiar de forma particular o en conjunto la logística urbana, el transporte de personas y los conductores individuales. En este caso los *datastreams* son generados de forma distribuida por cada agente involucrado y pueden ser procesados de forma centralizada o distribuida según el esquema elegido y los recursos disponibles.

Partiendo del objetivo general de lograr la Optimización de procesos de negocio por recomendación, surgen dos líneas de investigación. La primera propone la composición de modelos de analítica prescriptiva para superar los inconvenientes descritos. Estos modelos serán parte esencial de un proceso de mejora continua basado en la recomendación de acciones operativas y tácticas a fin de mantener el rendimiento del sistema en los valores deseados.

La segunda línea, está relacionada con el aprendizaje supervisado. Los algoritmos de aprendizaje automático pueden resolver de manera eficiente problemas con grandes conjuntos de datos complejos de distintas fuentes: web, redes sociales, eventos, teléfonos, telescopios, imágenes satelitales, entre otros. El algoritmo k-NN es muy conocido y ampliamente utilizado por su sencillez y robustez, descrito en 1967 por Cover y Hart [5]. Este método se basa en la idea intuitiva de que objetos similares pertenecen a la misma clase, entonces la clase de un objeto puede ser inferida a partir de la clase a la que pertenecen los objetos (o el objeto) más parecidos de la muestra de aprendizaje. La idea de similitud se refleja formalmente en el concepto de distancia.

La implementación más simple de k-NN tiene como gran inconveniente determinar los k-vecinos más cercanos de un objeto. Para esto se

requiere calcular las distancias entre él y todos los demás objetos de entrenamiento. Con el tamaño creciente de los conjuntos de datos, este enfoque se vuelve ineficiente, resultando en una búsqueda de complejidad lineal. Se han desarrollado variantes de k-NN, las cuales no sólo consideran la forma en que se comparan los objetos, sino también en las técnicas de programación, como es el caso de HPC, en particular GPU.

Para la primera línea, se propone el uso de agentes de aprendizaje por refuerzo [29], junto a técnicas provenientes de la Inteligencia Computacional: redes neuronales como modelos, teoría de conjuntos difusos como lenguaje de especificación, y métodos numéricos y metaheurísticos para el entrenamiento de tales modelos [7][28][33]. Para la segunda, el objetivo es acelerar y mejorar los métodos k-NN desarrollados considerando su utilización en problemas con grandes volúmenes de datos.

En ambos casos, se considera satisfacer la necesidad de dar rápida respuesta a los procesos de negocios dinámicos, en consecuencia es mandatorio pensar en la aplicación de modelos/paradigmas de HPC [9][15][22] portables, particularmente en GPU.

3. RESULTADOS OBTENIDOS/ESPERADOS

Por un lado, se encuentra en desarrollo un agente de aprendizaje por refuerzo para un caso de enrutamiento de vehículos, con suministro de información y re-encaminamiento en tiempo real. Este caso es uno de logística urbana. Se espera completar el desarrollo de un modelo prescriptivo dirigido por los datos, basado en la recomendación automática y proactiva de acciones operativas y tácticas destinadas a mantener los indicadores de rendimiento del sistema dentro de los valores deseados. Por el otro, se está trabajando en la adecuación de las implementaciones paralelas de k-NN considerando múltiples GPU.

Como uno de los objetivos es lograr soluciones paralelas portables, de costo predecible, capaz de explotar las ventajas de modernos ambientes HPC a través de

herramientas y “frameworks de computación” de alto nivel [15][22], los primeros desarrollos se están haciendo en GPU Nvidia con CUDA [18], y se prevé el uso de otras tecnologías como OpenCL[4] y OneAPI[20].

4. FORMACIÓN DE RECURSOS HUMANOS

Los resultados esperados respecto a la formación de recursos humanos son el desarrollo de 1 tesis de doctorado, 2 de maestría y de varias tesinas de grado en las universidades intervinientes.

5. BIBLIOGRAFÍA

- [1] M.Barrionuevo, M.Lopresti, N.Miranda, and M.F. Piccoli. “Solving a big-data problem with gpu: the network traffic analysis”. *Journal of Comp. Sc. and Techn.*, 15(01): Pp.30–39, Apr. 2015.
- [2] A. Bifet and J. Read. “Ubiquitous artificial intelligence and dynamic data streams”. In *Proceedings of the 12th ACM International Conference on Distributed and Event-Based Systems, DEBS '18*, Pp. 1–6, Association for Computing Machinery. New York, USA, 2018.
- [3] K. Braekers, K. Ramaekers and I. Van Nieuwenhuysse, "The vehicle routing problem: State of the art classification and review", *Comp. & Ind. Eng.*, vol. 99, pp. 300-313, 2016.
- [4] M. Breyer, G. Daiss and D. Pflüger. “*Performance-Portable Distributed k-Nearest Neighbors Using Locality-Sensitive Hashing and SYCL*”. *Int. Workshop on OpenCL*. ISBN 9781450390330. Association for Computing Machinery. 2021
- [5] T. Cover and P. Hart. 1967. *Nearest neighbor pattern classification*. *IEEE Transactions on Information Theory* 13, 1, 7. Jan. 1967. doi.org/10.1109/tit.1967.1053964.
- [6] G. Deka. "Big data predictive and prescriptive analytics". In *Handbook of research on cloud infrastructures for Big Data analytics*. IGI Global. p. 370-391. 2014.
- [7] A. Ebrahimnejad and J. L. Verdegay. “Fuzzy sets-based methods and techniques for modern analytics”. Springer Int. Publishing. 2018.
- [8] O. Etzion and P. Niblett, "Event processing in action", Manning, 2011.
- [9] C. Gröger, H. Schwarz, and B. Mitschang. “Prescriptive analytics for recommendation-based business process optimization”. In *Int. Conf. on Business Information Systems*, pages 25–37. Springer, 2014.
- [10] A. Hinze, K. Sachs and A. Buchmann. "Event-based applications and enabling technologies". *Proc. Third ACM Int. Conf. on Dist. Event-Based Systems*. pp. 1-15. 2009.
- [11] C. Holsapple, A. Lee-Post and R. Pakath. “A unified foundation for business analytics. *Decision Support Systems*”, 64:130{141, 2014.
- [12] M. Kaur Saggi and S. Jain. “A survey towards an integration of big data analytics to big insights for value-creation”. *Inf. Proc.& Management*, 54(5):758-790, 2018.
- [13] J. Keller, D. Liu and D. Fogel. "Fundamentals of computational intelligence: neural networks, fuzzy systems, and evolutionary computation". Wiley & Sons. 2016.
- [14] T. Kolajo, D. Olawande and A. Ayodele. "Big data stream analysis: a systematic literature review". *Journal of Big Data*, vol. 6, no 1, pp. 1-30, 2019.
- [15] S. Kurgalin and S. Borzunov, “A Practical Approach to High-Performance Computing”. Springer. 2019.
- [16] S. Melo, J. Macedo and P. Baptista, "Guiding cities to pursue a smart mobility paradigm: An example from vehicle routing guidance and its traffic and operational effects". *Research in Transp. Economics*, vol. 65, p. 24-33, 2017.
- [17] M. Minelli, M. Chambers, and A. Dhiraj. “Big data, big analytics: emerging business intelligence and analytic trends for today’s businesses”. Vol. 578. Wiley & Sons, 2013.
- [18] Nvidia. “*CUDA C++ Programming Guide, Design Guide*”.

- https://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf. 2019.
- [19] T. Olanrewaju. "The rise of the digital "bank."<https://www.mckinsey.com/business-functions/digital-mckinsey/ourinsights/the-rise-of-the-digital-bank/>. Accessed: 2020.
- [20] OneAPI Specification, release 1.0-rev 3. <https://spec.oneapi.com/versions/1.0-rev-3/>. 2020.
- [21] T. J. Owens, "Survey of event processing", Air Force Research Lab Rome NY Inf. Directorate, 2007.
- [22] P. Pacheco. "An Introduction to Parallel Programming", 1Ed., San Francisco, USA: Morgan Kaufmann Publishers Inc., 2011.
- [23] F. Pigni, G. Piccoli, and R. Watson. "Digital data streams: Creating value from the real-time flow of big data". *California Management Review*, 58(3):5–25, 2016.
- [24] S. Ramírez-Gallego, B. Krawczyk, S. García, M. Wozniak and F. Herrera. "A survey on data preprocessing for data stream mining: Current status and future directions". *Neurocomputing*, 239:39 – 57, 2017.
- [25] S. Russell and P. Norvig. "Inteligencia Artificial: un enfoque moderno". 2004.
- [26] R. Sharda, D. Delen, and E. Turban. "Business Intelligence and Analytics". 2015.
- [27] R. Sharda, D. Delen, and E. Turban. "Analytics, Data Science, & Artificial Intelligence". 2020.
- [28] N. Siddique and H. Adeli. "Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing". Wiley & Sons, 2013.
- [29] R. Sutton and A. Barto. "Reinforcement learning: An introduction". MIT press. 2018.
- [30] U. Thirathon, B. Wieder, Z. Matolcsy and M.L. Ossimitz. "Impact of big data analytics on decision making and performance". In *Int. Conf. on Enterprise Systems, Accounting and Logistics*, 2017.
- [31] M. Turner, D. Budgen and P. Brereton, "Turning software into a service". In *Computer*, vol. 36, no. 10, pp. 38-44, Oct. 2003.
- [32] C. Wrench, et al. "Data stream mining of event and complex event streams: A survey of existing and future technologies and applications in big data." *Enterprise Big Data Engineering, Analytics, and Management*. IGI Global, 2016. Pp. 24-47.
- [33] L.A. Zadeh. "Fuzzy logic, neural networks, and soft computing". *Comm. ACM* 37,3, 77–84. <https://doi.org/10.1145/175247.175255>. March 1994.