

BIGDATA EN LA EDUCACIÓN

Autores: Doc. Ing. Rodolfo Bollatti (Director Investigación)

Universidad Abierta Interamericana – Buenos Aires
Tecnología Informática
Secretaría de Investigación – CAETI

E-mail: Rodolfo.Bollatti@UAI.edu.ar

Telefono: 011-1536584700

Resumen

BigData se está convirtiendo en todos los entornos informatizados en el gran tema del momento, también lo va a ser en la educación. Con “BigData” nos referimos al hecho de que la producción y almacenamiento de información on-line está creciendo de forma tan desmesurada que las soluciones que en general han gestionado esta información a nivel corporativo, ya no son suficientes. Así como los profesionales y los modelos que tienen que analizar estos datos y convertirlos en elementos que resulten útiles para la toma de decisiones

El objetivo general de este proyecto es revelar la tendencia actual y futura sobre modelos basados en bigdata que ayuden a las instituciones educativas a identificar estudiantes en riesgo, de forma de intervenir con el fin de reducir la deserción y aumentar las tasas de graduación de los alumnos.

Palabras clave: BigData, Academic Analytics.

Contexto

El proyecto se basa en el desarrollo de conocimiento sobre la tecnología innovadora del BigData y el análisis de los modelos necesarios sobre dicha tecnología que aportan valor agregado a las instituciones educativas. Desarrollado desde la facultad de Tecnología de Información de la Universidad Abierta Interamericana, y la Secretaría de Investigación de UAI (CAETI)

Introducción

Por Big data nos referimos exactamente al tratamiento y análisis de enormes repositorios de datos, tan desproporcionadamente grandes y desestructurados que resulta imposible tratarlos con las herramientas de bases de datos y analíticas convencionales. La tendencia se encuadra en un entorno que no nos suena para nada extraño: la proliferación de páginas web, aplicaciones de imagen y vídeo, redes sociales, dispositivos móviles, aplicaciones, sensores, *internet de las cosas*, etc. capaces de generar, según IBM, más de 2.5 quintillones de bytes al día, hasta el punto de que el 90% de los datos del mundo han sido creados durante los últimos dos años. Hablamos de un entorno absolutamente relevante para muchos aspectos, desde el análisis de fenómenos naturales como el clima o de datos sismográficos, hasta entornos como salud, seguridad o, por supuesto, el ámbito educativo.

Las características del bigdata se centran en las tres v:

Volumen: Grandes volúmenes de datos, a partir de TeraBytes o incluso PetaBytes.

Variación: El concepto de Big Data también suele venir acompañado de diversos tipos de fuentes de datos, ya sean estructurados o no estructurados.

Velocidad: La frecuencia de las actualizaciones de estas grandes bases de datos también es un punto muy a tener en cuenta. Es por ello por lo que su procesamiento y posterior análisis también ha de realizarse prácticamente en tiempo real para poder mejorar la toma de decisiones en base a la información generada

Bigdata permite procesar gigantescos volumen de datos, no por medio de una supercomputadora, sino a través de computadoras pequeñas unidas en redes de procesamiento. Utilizaremos tecnología Hadoop (código abierto y gratuito) la cual permite a las aplicaciones trabajar con miles de nodos y petabytes de datos. Hadoop se inspiró en documentos Google.

Las soluciones de bigdata para en ámbito académico se pueden centrar en el aspecto institucional (Academic Analytics) como en el aprendizaje (Learning Analytics). En este proyecto nos centraremos en el aspecto institucional, relevando y desarrollando modelos

que permitirán por medio del BigData optimizar la deserción en el ámbito universitario como así también realizar un seguimiento más exhaustivo de los alumnos.

Nuestro alcance propone utilizar la información disponible (redes sociales, sistemas educativos, webs institucionales, etc) con el objetivo de generar e implementar modelos que permitan cumplir con nuestros objetivos planteados.

También abordaremos las problemáticas éticas que acompañan el desarrollo de dichos avances tecnológicos para el ámbito académico, la privacidad y la manipulación de datos sensibles ponen en evidencia los riesgos de tales prácticas.

Resultados y Objetivos

En la actualidad el proyecto se encuentra en su primera fase de maduración y análisis detallado.

A continuación se describen los correspondientes objetivos específicos a desarrollar:

- a) Generación de un entorno de desarrollo sobre tecnología BigData (Hadoop y plataforma de análisis).
- b) Estudio de modelos actuales y desarrollo de modelos específicos, que permitan identificar estudiantes en riesgo, de forma de tomar decisiones tendientes a intervenir con el fin de reducir la deserción y aumentar las tasas de graduación.
- c) Estudio de riesgos y cuestiones éticas.
- d) Desarrollo de un caso práctico sobre una institución académica, análisis de resultados y objeciones.

Formación de Recursos Humanos

A continuación se describen los productos asociados a dicho proyecto:

| Producto | Nº |
|--|----|
| Artículos en Revistas Internacionales con referato | 1 |
| Artículos en Revistas nacionales con referato | 2 |
| Presentaciones en Congresos | 3 |
| Tesis aprobadas | 1 |
| Recursos Humanos formados | 5 |

Referencias

- Campbell, J.P. DeBlois, P.B., & Oblinger, D.G. (2007). Academic analytics: A new tool for a new era.

- Carmean, C., & Mizzi, P. (2010). The case for nudge analytics.

- Norris, D., Baer, L., Leonard, J., Pugliese, L., & Lefrere, P. (2008). Action analytics: Measuring and improving performance that matters in higher education.

- Sarah Lacy, "Peter Thiel: We're in a Bubble and It's Not the Internet, It's Higher Education,"

- Bollier, D. (2010) 'The Promise and Peril of Big Data'

- Berry, D. (2011) 'The Computational Turn: Thinking About the Digital Humanities', Culture Machine.

- Boyd, d. and Ellison, N. (2007) 'Social Network Sites: Definition, History, and Scholarship',

- Acquisti, A. & Gross, R. (2009) 'Predicting Social Security Numbers from Public Data'