|  |  |  |  |
| --- | --- | --- | --- |
|  | | **FACULTAD DE INGENIERÍA**  **MAESTRÍA EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN** | |
| **TRABAJO DE GRADO - PREINSCRIPCIÓN PROYECTO DE APLICACIÓN** | | | |
| **TÍTULO DEL PROYECTO** | **Big Data Aplicada a Health Care** | | |
| **NOMBRE**  **ESTUDIANTE** | |  | | --- | | **Wilson Alzate Calderón** | | CC 1.32.400.001 | | **CORREO**  **ELECTRÓNICO** | [walzate@javeriana.edu.co](mailto:juanpp@javeriana.edu.co)  [wilson.alzate@gmail.com](mailto:juanito@gmail.com) |
| **DIRECTOR DE TRABAJO DE GRADO** | Ing. Alexandra Pomares PhD  [pomares@javeriana.edu.co](mailto:egon@javeriana.edu.co)  Universidad Javeriana  Profesor Ing. Sistemas | **MODALIDAD MISyC** | Profundización |
| **GRUPO Y LÍNEA DE INVESTIGACIÓN** | ISTAR - Sistemas de Información e Ingeniería de Software  Sub-línea - Gestión de Información |

|  |  |
| --- | --- |
| **PROBLEMA**  **DE**  **APLICACIÓN** | En el 2014 el monto de información almacenada en el mundo se estima que estará alrededor de 1200 exabytes, de los cuales menos del 2% es no digital[1]. El 80% de los datos relacionados a los negocios existen en un formato no estructurado[4]. Si este monto fueran solamente libros impresos, ellos cubrirían la superficie entera de Estados unidos en 52 capas de grueso. Google procesa más de 24 Petabytes de datos por día y Facebook obtiene más de 10 millones de fotos cada hora[1].  Teniendo en cuenta estas grades cantidades de datos que se están generando actualmente, surge el paradigma de Big Data, que es la habilidad de la sociedad de aprovechar la información de nuevas formas para producir bienes y servicios que generen valor. Siendo además la forma de aplicar matemáticas a grandes cantidades de datos con el fin de inferir probabilidades[1]. Hay que tener además en cuenta que Big Data se compone de cuatro dimensiones [5]:   * Volumen (Escala de los datos): 6 millones de personas usando celulares, 2.5 Quintillones de Bytes creados diariamente. * Variedad (Diferentes formas de datos): Para el 2014 se estima que van a haber 420 millones de dispositivos de monitoreo de salud portables, 4.2 billones de horas de video son vistas en YouTube cada mes. * Velocidad (Análisis de flujo de datos): Los automóviles modernos tienen cerca de 100 sensores, la bolsa de valores de Nueva York captura un Terabyte de información durante cada sesión de negociación. * Veracidad (Incertidumbre de los datos): Uno de cada tres líderes de negocios no creen completamente en la información que usan para tomar decisiones. La baja calidad en los datos le cuesta a la economía de los Estados Unidos un trillón de dólares al año.   En el contexto del cuidado de la salud, esta tecnología se puede aplicar al diagnóstico de enfermedades, recomendando tratamientos. Un ejemplo de esto es si millones de historias clínicas electrónicas revelan que los enfermos de cáncer que toman cierta combinación de aspirina y jugo de naranja ven mejoría en su enfermedad, la causa exacta de dicha mejoría puede ser menos importante que el hecho que ellos hayan vivido más tiempo[1].  Una de las muchas promesas de la ciencia de datos es que si se recolectan suficientes datos acerca de los tratamientos médicos y se usan de manera efectiva, será posible predecir de una mejor manera qué tratamientos serán más efectivos para cada paciente y cuáles no[3].  Se plantea entonces la posibilidad de generar un modelo de aplicación de Big Data en el domino del Cuidado de la salud, extendiéndolo a un caso de estudio, previendo problemáticas como[4]:   * Análisis de datos no estructurados contenidos en las historias clínicas electrónicas usando técnicas computacionales. * El análisis de las historias clínicas electrónicas para producir tableros de mando (Dashboards) que guíen las decisiones médicas. * Convertir iniciativas médicas personalizadas en prácticas clínicas * Entrega de información directamente a los pacientes, empoderándolos a jugar rol mucho más activo. * Crear un perfil clínico personalizado basado en filtro colaborativo[2] |
| **ASIGNATURAS MISyC APROBADAS**  **RELACIONADAS CON EL PROYECTO** | |  |  | | --- | --- | | Sistemas Inteligentes. | Diseño y Evaluación de Arquitecturas de Software. | | Arquitectura Empresarial | Gestión de Proyectos Informáticos | | Modelos Avanzados de Optimización | Diseño de Software Basado en Patrones | |
| **BIBLIOGRAFÍA** | [1]V. Mayer-Schönberger and K. Cukier, Big data: a revolution that will transform how we live, work, and think. Boston: Houghton Mifflin Harcourt, 2013.  [2]N. V. Chawla and D. A. Davis, “Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework,” Journal of General Internal Medicine, vol. 28, no. S3, pp. 660–665, Jun. 2013.  [3]T. O’Reilly, M. Loukides, J. Steele, and C. Hill, How Data Science Is Transforming Health Care, 1st ed. O’Reilly Media, 2012.  [4]T. B. Murdoch, “The Inevitable Application of Big Data to Health Care,” JAMA, vol. 309, no. 13, p. 1351, Apr. 2013.  [5]“IBM Big Data – What is Big Data – United States,” 17-Nov-2013. [Online]. Available: http://www.ibm.com/big-data/us/en/. [Accessed: 22-Nov-2013]. |