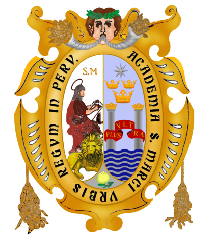
**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

(Universidad del Perú, DECANA DE AMÉRICA)

**FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA**

**EAP: INGENIERÍA DE SISTEMAS**

******

**SISTEMA WEB PARA DETERMINAR LOS PERFILES DE LA DEMANDA LABORAL DE LOS PROFESIONALES DE**

**INGENIERÍA DE SISTEMAS E INGENIERÍA DE SOFTWARE MEDIANTE WEB SCRAPING EN LOS PORTALES**

**LABORALES PARA LA TOMA DE DECISIONES**

**POR:**

MORÁN BAYONA, Pablo Misael

**ASESOR**

CASTRO LEÓN, Gloria Helena

**Lima – Perú**

2017

**Capítulo 1: INTRODUCCIÓN**

* 1. **Antecedentes**

Los portales laborales se han convertido en una valiosa fuente de información. Inicialmente tenían como público objetivo a los empleadores que buscaban recursos humanos con determinados perfiles para cubrir las plazas en las empresas y por otro lado a los profesionales en búsqueda de nuevas oportunidades laborales. Actualmente se cuenta con plataformas laborales que agrupan a un conjunto de portales laborales que referencian a una diversidad de oportunidades laborales disponibles en el mundo, es así como se ha identificado una excelente oportunidad para el aprovechamiento de esta rica fuente de información disponible en internet para otros rubros de aplicación como lo es el estudio de la demanda social como guía de referencia para la actualización de planes curriculares de las carreras profesionales.

Tenemos el caso de Mansourvar M. (2011) quien realizó una investigación en la Universidad de Malaya (UM) para obtener el grado de magister con el propósito de permitir al estudiante de la universidad, a través de la creación de un portal, poder elegir qué cursos seguir para su formación profesional, así como también poder elegir su especialidad tomando en cuenta la información del mercado laboral de modo que al finalizar su carrera pudieran conseguir un empleo con facilidad, que era algo alejado de la realidad, ya que en Malasia el 70% de egresados con estudios superiores no conseguía empleo al terminar sus estudios. El portal fue desarrollado y luego utilizado por usuarios finales satisfactoriamente cumpliendo con los objetivos propuestos en la investigación.

Existen casos donde la extracción de información de páginas web a través de la técnica Web Scraping fue de mucha utilidad en la recolección de información para su posterior análisis y utilización.

Casos como el de Blobel M. (2015) que desarrolló una investigación en la Universidad de Islandia (Háskóli Íslands) para obtener el grado de magister con el propósito de intentar reconstruir y categorizar los manuscritos ubicados en https://handrit.is/ escritos en Old Norse (lenguaje alemán antiguo) utilizando Web Scraping para la obtención de información y conectando los manuscritos a través de nodos y aplicando una metodología de análisis de redes para reunirlos en grupos con contenido similar y de esta forma identificar a los de “género étnico”, es decir, los que tienen contenido distinto. La información obtenida a través de Web Scraping probó ser de utilidad para la exploración con la metodología de análisis de redes.

Actualmente existe una necesidad de que las universidades tengan alineados los cursos que imparten con la necesidad del mercado laboral y que el número de plazas para cada carrera en las universidades sean las que se necesitan en la sociedad. De este modo se evita que se produzca con el tiempo un exceso de profesionales que supera en gran medida a lo que el mercado espera. Por otro lado se desea crear a los profesionales con los conocimientos que el mercado necesita para que sean los más competentes en el rubro que elijan.

* 1. **Definición del Problema**

Las universidades no cuentan con una investigación o alguna herramienta de donde apoyarse al momento seleccionar los cursos o materias que se impartirán en su plan de estudio para las carreras de Ingeniería de Sistemas e Ingeniería de Software.

De la misma manera, no se cuenta con un indicador confiable y basado en una investigación para que el número de plazas ofrecidas en las carreras de Ingeniería de Sistemas e Ingeniería de Software vayan acorde con las necesidades del mercado laboral.

* 1. **Objetivos**
     1. **Objetivo General**

Desarrollar un Sistema Web que permita determinar los perfiles de la demanda laboral en Latinoamérica de los profesionales de Ingeniería de Sistemas e Ingeniería de Software mediante el uso de la técnica Web Scraping en los portales laborales.

* + 1. **Objetivos Específicos**
* Identificar los portales laborales que se utilizarán para la extracción de información
* Formular un modelo a partir de la data extraída para su clasificación.
* Desarrollar un sistema web que exprese las funcionalidades de la propuesta.
  1. **Justificación**

La presente investigación se enfoca primero en la necesidad de las carreras profesionales de Ingeniería de Sistemas e Ingeniería de Software de lograr su acreditación; El modelo de acreditación considera en los estándares dos y cinco la importancia de la demanda social como la fuente de información de valor en la definición de la oferta académica en cuanto a su pertinencia y tamaño. La pertinencia de la formación que se brinda al estudiante con respecto al perfil de egreso y que esta debe responder a las expectativas de los grupos de interés.

Los cursos que se imparten deben estar alineados con las necesidades del mercado laboral. Asimismo la proyección del número de plazas o vacantes para los procesos de admisión a las carreras profesionales de la universidad debe estar justificados con la demanda social de los perfiles laborales según especialidad para evitar un exceso o déficit de profesionales.

Esta investigación permitirá, mediante un sistema web, brindar información de valor de los perfiles de profesionales que se necesitan en el mercado laboral. La fuente de información base se obtendrá desde múltiples portales laborales utilizando la técnica Web Scraping y, mediante técnicas de clasificación, se expondrán los resultados del proceso analítico. De este modo se puede aportar a las universidades a tener un estudio de demanda social para la toma de decisiones concernientes a este tema.

**Lista de Referencias**

1. Blobel, M. (2015) "Scraping Parchment, Investigating Genre through Network Analysis of the Electronic Manuscript Catalogue Handrit.is" Tesis para obtener el grado de magister en ciencias de la computación.
2. CastrilloFernández, O. (2015) "Web Scraping, Applications and Tools", ePSIplatform Topic Report No. 2015 / 10 , December 2015
3. Khalil, S.; Fakir, M.(2017) "RCrawler: An R package for parallel web crawling and scraping", SoftwareX 6 (2017) 98–106.
4. Mansourvar, M. (2011) "Development Of A Job Web Portal To Capture Industry’s Needs", Tesis para obtener el grado de magister en ciencias de la computación.
5. Maureen S. (2014) "Erecruitment: A comparison with traditional recruitment and the influences of social media A qualitative and quantitative review", Tesis para obtener el grado de Bachiller en Administración de Negocios.
6. Mytna, L.; Beblavy, M.; Thum, A.(2014), "Using Internet Data to Analyse the Labour Market: A Methodological Enquiry", Germany.
7. Polidoro, F.; Giannini, R.; Lo Conte, R.; Mosca, S.; Rossetti, F. (2015) "Web scraping techniques to collect data on consumer electronics and airfares for Italian HICP compilation", Italy, Statistical Journal of the IAOS 31 (2015) 165–176