

PROYECTO DE INVESTIGACIÓN DATOS INFORMATIVOS

TIPO DE CONVOCATORIA

Proyecto Interno Proyecto Semilla Proyecto Junior Proyecto Multi e Interdisciplinario

Fecha de presentación (dd/mm/aa): 06/09//2017

Título del proyecto: Minería de datos, extracción de vector de características y obtención de modelos con reconocimiento de patrones y aprendizaje automático para detección de escenarios relacionados con el delito de tráfico de personas.

TIPOS DE INVESTIGACIÓN

Investigación básica

Investigación aplicada

DEPARTAMENTO(S) Y/O INSTITUCIÓN:

I. Departamento de Informática y Ciencias de la Computación

LÍNEA(S) DE INVESTIGACIÓN (verificable en el SAEW):

I. Machine Learning

RESUMEN DE INFORMACIÓN DEL DIRECTOR Y COLABORADORES

| Director | | | | |
|----------------------------------|---------------|-----|--|--|
| Apellidos y nombres | No. de Cédula | HSS | Departamento | Título de mayor nivel y mención. |
| Hernández Álvarez Myriam Beatriz | 170500930-4 | 8 | Departamento de Informática y ciencias de la Computación | Ph.D en Aplicaciones de la Informática |

| Codirector (Se aplica para todos los proyectos, el codirector será a su vez colaborador) | | | | |
|--|---------------|-----|--|------------------------------------|
| Apellidos y nombres | No. de Cédula | HSS | Departamento | Título de mayor nivel y mención. |
| Aguar Pontes Josafá de Jesus ✓ | 175694375-7 | 4 | Departamento de Informática y ciencias de la Computación | PhD. en Ciencias de la Computación |

| Colaborador | | | | |
|------------------------------|---------------|-----|--|-------------------------------------|
| Apellidos y nombres | No. de Cédula | HSS | Departamento | Título de mayor nivel y mención. |
| Burbano Acuña Marco Danilo ✓ | 171712059-4 | 8 | Departamento de Informática y ciencias de la Computación | M.Sc. en Ciencias de la Computación |

* HSS = Horas Semana Semestre

PROYECTO DE INVESTIGACIÓN

Proyecto Interno Proyecto Semilla Proyecto Junior Proyecto Multi e Inter Disciplinario

Investigación Básica

Investigación Aplicada

DEPARTAMENTO(S) Y/O INSTITUTOS:

1. Departamento de Informática y Ciencias de la Computación

LINEA(S) DE INVESTIGACIÓN:

1. Machine Learning

DISCIPLINA CIENTÍFICA (Marque X, solamente una opción)

| | |
|------------------------------|---|
| Ciencias Naturales y Exactas | |
| Ingeniería y Tecnologías | X |
| Ciencias Médicas | |
| Ciencias Agrícolas | |
| Ciencias Sociales | |
| Humanidades | |

OBJETIVO SOCIOECONÓMICO (Marque X, solamente una opción)

| | |
|---|---|
| Exploración y explotación del medio terrestre | |
| Ambiente | |
| Exploración y explotación del espacio | |
| Transporte, telecomunicaciones y otras infraestructuras | |
| Energía | |
| Producción y tecnología industrial | |
| Salud | |
| Agricultura | |
| Educación | |
| Cultura, ocio, religión y medios de comunicación | |
| Sistemas políticos y sociales, estructuras y procesos | |
| Defensa | |
| Avance general del conocimiento: I+D financiada con los Fondos Generales de Universidades (FGU) | X |
| Avance general del conocimiento: I+D financiados con otras fuentes | |



| |
|--|
| 1 Proyecto de Investigación |
| Título: Minería de datos, extracción de vector de características y obtención de modelos con reconocimiento de patrones y aprendizaje automático para detección de escenarios relacionados con el delito de tráfico de personas. |
| Resumen del proyecto (máximo 200 palabras) El presente proyecto tiene como objetivo utilizar mecanismos de inteligencia artificial para extraer información y obtener patrones referentes al delito de tráfico de personas en el Ecuador utilizando técnicas de minería de datos y reconocimiento de patrones. Los datos serán obtenidos utilizando varias fuentes tales como noticias en línea e información en redes sociales vinculadas con este tema. Mediante el uso de algoritmos de aprendizaje automático (machine learning) aplicados al procesamiento de lenguaje natural, se analizará, extraerán características y se clasificará la información obtenida con el objetivo de encontrar patrones y tendencias relacionadas al tráfico de personas. Dentro del proyecto se incluirá un pre-procesamiento léxico y sintáctico del corpus con el fin de prepararlo para una clasificación semántica. Se utilizarán herramientas como tablas de decisión y árboles para discriminar clases de atributos que ingresarán a algoritmos de aprendizaje automático parcialmente supervisado. A partir de información obtenida de las fuentes mencionadas, se creará un corpus que se procesará léxica y sintácticamente, se detectarán atributos relevantes y utilizando técnicas de aprendizaje automático se obtendrá una clasificación que permita modelar posibles escenarios vinculados con el delito de trata de personas. |
| Palabras clave (4-6): Tráfico de personas, aprendizaje automático, datos, extracción, análisis léxico, análisis sintáctico |

| |
|---|
| 2 Objetivos, limitaciones, hipótesis y resultados esperados de esta propuesta de investigación |
|---|

2.1 Objetivos

2.1.1 Objetivo General

- Generar un corpus con información referente al delito de tráfico de personas en Ecuador, procesarlo, extraer características y clasificarlo para obtener modelos de posibles escenarios relacionados con el mencionado tipo de crimen.

2.1.2 Objetivos Específicos

- a. Realizar minería de datos en periódicos en línea, redes sociales.
- b. Pre-procesar y etiquetar con resultados de análisis léxico y sintáctico.
- c. Definir formato del corpus de acuerdo con las especificaciones del análisis léxico y sintáctico.
- d. Implementar aprendizaje automático para establecer atributos relevantes para ingreso a clasificadores.
- e. Evaluar algoritmos de clasificación utilizados para obtener modelos de escenarios posibles.



- f. Presentar artículo científico para publicación en un congreso ISI Web of Knowledge o Scopus
- g. Presentar artículo científico para publicación en una revista ISI Web of Knowledge o Scopus

2.2 Limitaciones (Aspectos que quedan fuera del alcance del Proyecto de Investigación)

- a. Se aplicarán algoritmos ya desarrollados para aprendizaje automático.
- b. En la definición del alcance, se restringe la información a fuentes en línea que contengan datos en español.
- c. La extracción de características relacionadas con el crimen de tráfico de personas, será enfocado en el idioma español.

2.3 Hipótesis (Responden al problema de investigación)

- a. Es posible construir un corpus con información relevante al tráfico de personas usando como fuente noticias de periódicos en línea y datos de redes sociales.
- b. Usando fuentes en Internet es posible encontrar patrones relacionados al tráfico de personas.

2.3 Detalle de los resultados esperados (con relación a los objetivos)

- a. Publicaciones en ISI Web of Knowledge o Scopus que incluya información relacionada al delito de tráfico de personas en Ecuador, en formato adecuado para análisis léxico y sintáctico, pre-procesada, etiquetada y ordenada de acuerdo a relevancia en el tema y que además incluya un reporte del rendimiento de los algoritmos.
- b. Aplicación que genere el corpus con características relevantes y a partir de aprendizaje automático obtenga el modelo con los patrones que permitan identificar los escenarios relacionados con el delito de trata de personas.

| | |
|----------|--|
| 3 | Relevancia de la propuesta de investigación y su relación con la(s) líneas de investigación |
|----------|--|

El tráfico de personas es un crimen que consiste en una forma de esclavitud donde las víctimas son forzadas o actúan bajo coerción con fines de explotación. En el Ecuador se tiene la trata de personas como delito tipificado en el Código Orgánico Integral Penal (COIP) dentro de 20 artículos, que consideran ocho tipos de explotación: sexual, laboral, mendicidad, reclutamiento forzoso, adopciones ilegales, matrimonio servil y tráfico de órganos.

El número de víctimas en el mundo se estima que está entre 12 millones a 27 millones [1] y en Ecuador no se lo tiene completamente cuantificado, pero es un delito que genera ganancias millonarias. Para realizar estas actividades criminales se está extendiendo el uso de clasificados en línea y redes sociales. Los traficantes usan herramientas del siglo XXI, disponibles en Internet, como base para sus ilícitos como una forma de reclutamiento y publicidad para atraer víctimas. El rol de redes sociales y clasificados en línea en Ecuador no ha sido estudiado.

El trabajo que se plantea tiene como hipótesis que herramientas como minería de datos, mapeo y analítica de datos pueden servir para obtener patrones de textos en noticias en línea o conversaciones en redes sociales, que ayuden en la detección para prevención del tráfico de personas. Estas técnicas se estudiarán con fuentes obtenidas en los límites de la ciudad de Quito. Se analizará la relación entre el tráfico de personas y la información que se puede encontrar en Internet en periódicos en línea, Facebook y Twitter que pueden ser usados por traficantes para mercadear, reclutar, vender y explotar a las personas con propósitos criminales en forma explícita o encubierta.

Es importante mencionar que se tuvo un acercamiento con el Ministerio del Interior para tener acceso a información de primera mano en lo relacionado con datos criminales en Quito, sin embargo, por el



tema de transición de autoridades debido al nuevo gobierno, no se ha tenido aún una respuesta formal por parte de dicha institución para que proporcione una nueva fuente de información, por lo que no se puede asegurar que se vaya disponer de estos datos.

El presente proyecto, marca el inicio de una serie de proyectos de investigación requeridos para el desarrollo de una tesis doctoral en el marco de la segunda cohorte del programa doctoral en informática. Entre los aportes más relevantes de este proyecto está la ingeniería de características que se realizará para obtener los atributos que se ingresarán en los algoritmos de clasificación a partir del conjunto completo de variables que se encontrarán en las fuentes de datos. Con el vector de características se conformará un corpus para procesarlo con métodos de aprendizaje automático sintáctico – estructural y semántico para obtener los modelos o patrones que configuren escenarios de riesgo para la detección del delito de tráfico de personas. Estos procesos se realizarán en idioma español lo que constituye una característica especial del aporte del proyecto, tomando en cuenta que mucho del trabajo realizado en este campo específico se lo encuentra para datos en inglés.

4 Productos esperados

| Tipo de Producto: | Marcar con una "X" |
|---|--------------------|
| a. Publicaciones científicas (obligatorio); | X |
| b. Disertación a la comunidad politécnica; | |
| c. Trabajo de titulación de acuerdo a lo que establece el Reglamento de Régimen Académico y la Normativa Interna de la EPN; | |
| d. Aplicación tecnológica construida o implementada; | X |
| e. Patente presentada; | |
| f. Perfil de proyecto de mayor impacto científico, técnico, pedagógico o de innovación. | |

5 Descripción, metodología y diseño del proyecto

5.1 Descripción, metodología y diseño del proyecto (Máximo dos carillas)

En el presente proyecto se generará un corpus para procesamiento de información referente al tráfico de personas, para lo cual se realizará una extracción de datos de información en línea, tales como noticias de periódicos[2], [3] y redes sociales. Posteriormente se realizará un análisis de los datos extraídos con el fin de obtener información ordenada de acuerdo a la relevancia con respecto al tráfico de personas. Finalmente se realizará una evaluación de los algoritmos de aprendizaje automático utilizados.

La metodología a seguir para la consecución del presente proyecto, se resume en las siguientes actividades:

- Minería de datos de periódicos en línea.
- Minería de datos de redes sociales.
- Preprocesamiento de datos y extracción del vector de características relevantes para constituir el corpus.
- Análisis léxico, sintáctico.
- Etiquetado de los datos.
- Entrenamiento y evaluación de algoritmos de aprendizaje automático para obtener los modelos y patrones de escenarios de riesgo para el cometimiento del delito.

La metodología anteriormente detallada, se la definió en base a trabajos de investigación similares, cuyas referencias se presentan en el detalle de cada actividad a continuación:



Minería de datos de periódicos en línea y/o Dark Web

En la Dark Web, se puede encontrar diversas formas de publicidad en línea relacionadas al tráfico de personas, en páginas como backpage.com, adultfinder.com, cityvibe.com, eros.com, humaniplex.com, myredbook.com, sugardaddyforme.com [4], [5] existen anuncios camuflados que pueden ser extraídos y analizados. Para esta tarea se tiene planificado usar web spidering, que se refiere al proceso de explorar de forma intensa la Web [3], [6], es decir el Internet como tal, en otras palabras, una búsqueda automática de varias páginas web, las mismas que son utilizadas como base para la creación de un corpus en particular. Para el caso de este proyecto, lo que se plantea realizar mediante web spidering es obtener una lista de direcciones web (URLs) de las noticias relacionadas al tráfico de personas de un periódico en línea. Posteriormente se extraerá texto obtenido en esas direcciones.

Minería de datos de redes sociales

Los traficantes usan una variedad de medios para promocionar servicios ilícitos, incluyendo anuncios camuflados a través del Internet como clasificados y boletines en redes sociales asociados con servicios de masajes y damas de compañía [5], [7]. Considerando que el usuario de las redes sociales produce una variedad de contenido en varias plataformas sociales, la extracción de contenidos en estos medios convierte en una tarea crucial. En esta actividad se tiene planificado utilizar un API [8], [9] asociado a la red social a utilizar para la extracción de información y almacenarla en una base de datos para su posterior procesamiento [10]–[13].

Preprocesamiento de datos y formato de corpus

Esta tarea se refiere a todo el trabajo involucrado en la limpieza de los datos extraídos, como por ejemplo la conversión de HTML puro en una forma de texto que pueda ser utilizado como entrada para los algoritmos de procesamiento de lenguaje natural [6], [14]. Posteriormente, se identificará la información relevante para el tráfico de personas mediante la aplicación de varias técnicas de procesamiento como reemplazo de palabras negativas, palabras con mayor coincidencia de aparición, etc. [14].

Adicionalmente, este proceso se refiere a la definición de características [6] que identifican si un artículo está relacionado con el tráfico de personas, en otras palabras, definir las columnas con información trascendente al tráfico de personas en función de la base de conocimiento obtenida en las fases anteriores. De esta manera se espera tener definido un vector de características en el corpus para Facebook, Twitter y los periódicos en línea respectivamente.

Análisis léxico, sintáctico

Durante los últimos años se han desarrollado muchísimas herramientas para la descomposición léxica y sintáctica de los textos con el fin de procesarlos para obtener su significado semántico. En nuestro corpus pre-procesado se evaluarán algunas de estas herramientas para aplicarlas al corpus y obtener entradas enriquecidas con etiquetas que sirvan como entrada a los clasificadores que utilizarán algoritmos de aprendizaje automático parcialmente supervisado y supervisado.

Etiquetado de datos

La tarea de etiquetado de datos implica un procedimiento manual y/o automático [16], [17], donde se define si cada fila del conjunto de datos está altamente relacionada con el tráfico de personas o no [6]. El análisis léxico será automático usando herramientas como las disponibles en la Librería de Stanford. El análisis sintáctico podrá ser realizado en forma semiautomática.

Entrenamiento y evaluación de algoritmos de aprendizaje automático

Una vez definido el conjunto de datos con sus respectivas características, se entrenarán con varios algoritmos de aprendizaje automático. En este proyecto se tiene planificado la utilización de algoritmos de aprendizaje de máquina populares entre la comunidad de analítica de textos tales como: Support Vector Machines (SVM), Naive Bayes, k-Nearest Neighbors (k-NN) [15], para evaluar cada uno de los modelos generados e identificar aquel con el mejor resultado.

Referencias Bibliográficas:

- [1] UNODC, "INFORME MUNDIAL SOBRE LA TRATA DE PERSONAS."
- [2] M. A. Awal, J. Rabbi, S. I. Hossain, and M. M. A. Hashem, "Using linear regression to forecast future trends in crime of Bangladesh," in *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*, 2016, pp. 333–338.
- [3] R. Arulanandam, B. T. R. Savarimuthu, and M. A. Purvis, "Extracting crime information from online



- newspaper articles," *Proc. Second Australas. Web Conf. - Vol. 155*, pp. 31–38, 2014.
- [4] H. Alvari, P. Shakarian, and J. E. K. Snyder, "A non-parametric learning approach to identify online human trafficking," in *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*, 2016, pp. 133–138.
- [5] H. Wang, C. Cai, A. Philpot, M. Latonero, E. H. Hovy, and D. Metzler, "Data integration from open internet sources to combat sex trafficking of minors," in *Proceedings of the 13th Annual International Conference on Digital Government Research - dg.o '12*, 2012, p. 246.
- [6] V. Liu and J. Curran, "Web Text Corpus for Natural Language Processing.," *EACL*, 2006.
- [7] D. R. Silva, A. Philpot, A. Sundararajan, N. M. Bryan, and E. Hovy, "Data integration from open internet sources and network detection to combat underage sex trafficking," in *Proceedings of the 15th Annual International Conference on Digital Government Research - dg.o '14*, 2014, pp. 86–90.
- [8] M. K. Rahman, "DataViz: High velocity data visualization and retrieval of relevant information from social network," in *2015 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2015, pp. 1–7.
- [9] R. McCreddie, I. Soboroff, J. Lin, C. Macdonald, I. Ounis, and D. McCullough, "On building a reusable Twitter corpus," in *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '12*, 2012, p. 1113.
- [10] W. Ahmad and R. Ali, "Information retrieval from social networks: A survey," in *2016 3rd International Conference on Recent Advances in Information Technology (RAIT)*, 2016, pp. 631–635.
- [11] P. M. Zadeh, M. Mohi, and M. S. Moshkenani, "Mining social network for extracting topic of textual conversations," in *Proceedings of the 5th international conference on Soft computing as transdisciplinary science and technology - CSTST '08*, 2008, p. 232.
- [12] Xindong Wu, Xingquan Zhu, Gong-Qing Wu, and Wei Ding, "Data mining with big data," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 1, pp. 97–107, Jan. 2014.
- [13] F. Jiang, C. K. Leung, and A. G. M. Pazdor, "Big data mining of social networks for friend recommendation," in *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 2016, pp. 921–922.
- [14] W. Medhat, A. H. Yousef, and H. Korashy, "A framework of preparing corpora from social network sites for sentiment analysis," in *International Conference on Information Society (i-Society 2014)*, 2014, pp. 32–39.
- [15] R. J. Sethi, Y. Gil, H. Jo, and A. Philpot, "Large-scale multimedia content analysis using scientific workflows," in *Proceedings of the 21st ACM international conference on Multimedia - MM '13*, 2013, pp. 813–822.
- [16] R. Kapoor, M. Kejriwal, and P. Szekely, "Using contexts and constraints for improved geotagging of human trafficking webpages," in *Proceedings of the Fourth International ACM Workshop on Managing and Mining Enriched Geo-Spatial Data - GeoRich '17*, 2017, pp. 1–6.
- [17] M. Kejriwal and P. Szekely, "Information Extraction in Illicit Web Domains," in *Proceedings of the 26th International Conference on World Wide Web - WWW '17*, 2017, pp. 997–1006.

6 Infraestructura, equipos y fondos adicionales.

6.1 Infraestructura y equipos

La infraestructura necesaria para este proyecto está compuesta únicamente de los computadores personales que cada participante tiene asignado a su cargo, los mismos que se encuentran ubicados en la Facultad de Sistemas de la EPN.

6.2 Breve justificación del equipo requerido

La ejecución de los procedimientos para minería de datos, así como los algoritmos de aprendizaje automático, únicamente requieren de la ejecución de sus respectivos programas en un computador personal, el mismo que requiere tener instalados programas de acceso libre como Java, Python, Octave, etc. Dichos equipos se encuentran ubicados en la sala de profesores, al frente del departamento de informática y ciencias de la computación.

6.3 Fondos Adicionales

- No se cuenta con fondos adicionales



VICERRECTORADO DE INVESTIGACIÓN Y PROYECCIÓN SOCIAL

UNIDAD DE INVESTIGACIÓN

PRESUPUESTO PROYECTOS DE INVESTIGACIÓN



AÑO 1

| | |
|--------------------------------------|--|
| Director del proyecto | Título del proyecto |
| Dra. Myriam Hernández Alvarez | Minería de datos, extracción de vector de características y obtención de modelos con reconocimiento de patrones y aprendizaje automático para detección de escenarios relacionados con el delito de tráfico de personas. |

| Lista de Items | Cantidad | Unidad | Precio Unitario Referencial | Precio Total Referencial | Precio Unitario Referencial +Aporte IESS | Precio Total Referencial con IVA + Aporte del IESS |
|--|----------|--------|-------------------------------------|----------------------------------|--|--|
| 1 Contratación de servicios personales por contrato | | | | | | |
| 1.1 Ayudantes de investigación | 12 | mes | \$ 125.00 | \$ 1,500.00 | \$ 136.44 | \$ 1,637.25 |
| 1.2 Prestación de servicios profesionales (Homologado Escala de remuneración de servidores publicos) | | mes | \$ - | \$ - | \$ - | \$ - |
| Subtotal 1 | | | \$ 125.00 | \$ 1,500.00 | \$ 136.44 | \$ 1,637.25 |
| Lista de Items | Cantidad | Unidad | Precio Unitario Referencial sin IVA | Precio Total Referencial sin IVA | Precio Unitario Referencial con IVA | Precio Total Referencial con IVA |
| 2 Maquinaria equipos | | | | | | |
| 2.1 Item 1 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.2 Item 2 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.3 Item 3 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.4 Item 4 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.5 Item 5 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 2 | | | \$ - | \$ - | \$ - | \$ - |
| 3 Reactivos y materiales de laboratorio | | | | | | |
| 3.1 Item 1 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.2 Item 2 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.3 Item 3 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.4 Item 4 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.5 Item 5 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 3 | | | \$ - | \$ - | \$ - | \$ - |
| 4 Literatura especializada | | | | | | |
| 4.1 Item 1 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.2 Item 2 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.3 Item 3 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.4 Item 4 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.5 Item 5 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 4 | | | \$ - | \$ - | \$ - | \$ - |
| 5 Viajes técnicos y de muestreo | | | | | | |
| 5.1 Pasajes al interior | | | \$ - | \$ - | \$ - | \$ - |
| 5.2 Viaticos al interior | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 5 | | | \$ - | \$ - | \$ - | \$ - |
| 6 Presentación de ponencias en congresos internacionales y publicaciones | | | | | | |
| 6.1 Pasajes al exterior | 1 | | \$ 2,500.00 | \$ 2,500.00 | \$ 2,800.00 | \$ 2,800.00 |
| 6.2 Viaticos al exterior | 1 | | \$ 2,000.00 | \$ 2,000.00 | \$ 2,240.00 | \$ 2,240.00 |
| 6.3 Pago de inscripción y publicaciones | 1 | | \$ 2,000.00 | \$ 2,000.00 | \$ 2,240.00 | \$ 2,240.00 |
| Subtotal 6 | | | \$ 6,500.00 | \$ 6,500.00 | \$ 7,280.00 | \$ 7,280.00 |
| TOTAL | | | | \$ 8,000.00 | | \$ 8,917.25 |



VICERRECTORADO DE INVESTIGACIÓN Y PROYECCIÓN SOCIAL

UNIDAD DE INVESTIGACIÓN

PRESUPUESTO PROYECTOS DE INVESTIGACIÓN



AÑO 2

| | |
|--------------------------------------|--|
| Director del proyecto | Título del proyecto |
| Dra. Myriam Hernández Alvarez | Minería de datos, extracción de vector de características y obtención de modelos con reconocimiento de patrones y aprendizaje automático para detección de escenarios relacionados con el delito de tráfico de personas. |

| Lista de Items | Cantidad | Unidad | Precio Unitario Referencial | Precio Total Referencial | Precio Unitario Referencial - Aporte IESS | Precio Total Referencial con IVA + Aporte del IESS |
|--|----------|--------|-------------------------------------|----------------------------------|---|--|
| 1 Contratación de servicios personales por contrato | | | | | | |
| 1.1 Ayudantes de Investigación | 6 | mes | \$ 125.00 | \$ 750.00 | \$ 136.44 | \$ 818.63 |
| 1.2 Prestación de servicios profesionales (Homologado Escala de remuneración de servidores publicos) | | mes | \$ - | \$ - | \$ - | \$ - |
| Subtotal 1 | | | \$ 125.00 | \$ 750.00 | \$ 136.44 | \$ 818.63 |
| Lista de Items | Cantidad | Unidad | Precio Unitario Referencial sin IVA | Precio Total Referencial sin IVA | Precio Unitario Referencial con IVA | Precio Total Referencial con IVA |
| 2 Maquinaria equipos | | | | | | |
| 2.1 Item 1 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.2 Item 2 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.3 Item 3 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.4 Item 4 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| 2.5 Item 5 (Detallar nombre de la maquinaria y equipos solicitado) | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 2 | | | \$ - | \$ - | \$ - | \$ - |
| 3 Reactivos y materiales de laboratorio | | | | | | |
| 3.1 Item 1 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.2 Item 2 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.3 Item 3 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.4 Item 4 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| 3.5 Item 5 (Detallar nombre de los insumos y reactivos) | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 3 | | | \$ - | \$ - | \$ - | \$ - |
| 4 Literatura especializada | | | | | | |
| 4.1 Item 1 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.2 Item 2 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.3 Item 3 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.4 Item 4 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| 4.5 Item 5 (Detallar nombre del libro) | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 4 | | | \$ - | \$ - | \$ - | \$ - |
| 5 Viajes técnicos y de muestreo | | | | | | |
| 5.1 Pasajes al interior | | | \$ - | \$ - | \$ - | \$ - |
| 5.2 Viáticos al interior | | | \$ - | \$ - | \$ - | \$ - |
| Subtotal 5 | | | \$ - | \$ - | \$ - | \$ - |
| 6 Presentación de ponencias en congresos internacionales y publicaciones | | | | | | |
| 6.1 Pasajes al exterior | | | | | | |
| 6.2 Viáticos al exterior | | | | | | |
| 6.3 Pago de inscripción y publicaciones | | | | | | |
| Subtotal 6 | | | \$ - | \$ - | \$ - | \$ - |
| TOTAL | | | | \$ 750.00 | | \$ 818.63 |

**VICERRECTORADO DE INVESTIGACIÓN Y PROYECCIÓN SOCIAL**

UNIDAD DE INVESTIGACIÓN

PRESUPUESTO PROYECTOS DE INVESTIGACIÓN



| Director del proyecto | Título del proyecto |
|-------------------------------|--|
| Dra. Myriam Hernández Álvarez | Minería de datos, extracción de vector de características y obtención de modelos con reconocimiento de patrones y aprendizaje automático para detección de escenarios relacionados con el delito de tráfico de personas. |

Presupuesto consolidado sin IVA

| AÑO | Contratación de servicios personales por contrato | Maquinaria y equipo | Reactivos y materiales de laboratorio | Literatura especializada | Viajes técnicos y de muestreo | Presentación de ponencias en congresos internacionales y publicaciones | Total sin IVA |
|--------------|---|---------------------|---------------------------------------|--------------------------|-------------------------------|--|--------------------|
| 1 | \$ 1,500.00 | \$ - | \$ - | \$ - | \$ - | \$ 6,500.00 | \$ 8,000.00 |
| 2 | \$ 750.00 | \$ - | \$ - | \$ - | \$ - | \$ - | \$ 750.00 |
| 3 | \$ - | \$ - | \$ - | \$ - | \$ - | \$ - | \$ - |
| TOTAL | \$ 2,250.00 | \$ - | \$ - | \$ - | \$ - | \$ 6,500.00 | \$ 8,750.00 |

Presupuesto consolidado con IVA

| AÑO | Contratación de servicios personales por contrato | Maquinaria y equipo | Reactivos y materiales de laboratorio | Literatura especializada | Viajes técnicos y de muestreo | Presentación de ponencias en congresos internacionales y publicaciones | Total con IVA |
|--------------|---|---------------------|---------------------------------------|--------------------------|-------------------------------|--|--------------------|
| 1 | \$ 1,637.25 | \$ - | \$ - | \$ - | \$ - | \$ 7,280.00 | \$ 8,917.25 |
| 2 | \$ 818.63 | \$ - | \$ - | \$ - | \$ - | \$ - | \$ 818.63 |
| 3 | \$ - | \$ - | \$ - | \$ - | \$ - | \$ - | \$ - |
| TOTAL | \$ 2,455.88 | \$ - | \$ - | \$ - | \$ - | \$ 7,280.00 | \$ 9,735.88 |

DECLARACIÓN FINAL

TIPO DE PROYECTO

Proyecto Interno Proyecto Semilla Proyecto Junior Proyecto Multi e Interdisciplinario

TIPO DE INVESTIGACIÓN

Investigación básica

Investigación aplicada

TÍTULO DEL PROYECTO

Minería de datos, extracción de vector de características y obtención de modelos con reconocimiento de patrones y aprendizaje automático para detección de escenarios relacionados con el delito de tráfico de personas.

DECLARACIÓN DEL DIRECTOR DEL PROYECTO

El equipo de investigadores, representado por el Director del Proyecto declara lo siguiente:

- Que el presente proyecto es una obra original de este equipo de investigadores y por tanto, asumimos la completa responsabilidad legal en caso de que un tercero alegue la titularidad de los derechos intelectuales del proyecto, exonerando a la EPN de cualquier acción legal que se derive por esta causa.
- Que el presente proyecto no ha sido presentado en ninguna convocatoria de otra institución pública o privada solicitando el financiamiento total del presupuesto. El incumplimiento será causal para que el proyecto no sea tomado en consideración.
- Que, todos los bienes adquiridos en el proyecto permanecerán bajo la custodia y responsabilidad del director de proyecto.
- Que, aceptamos que si el proyecto genera algún producto o procedimiento susceptible de obtener de derechos de propiedad intelectual, de los cuales se deriven beneficios, estos serán compartidos entre los investigadores y las instituciones participantes en el proyecto.

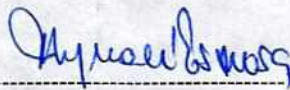


Firma del Director del Proyecto
Nombre: Myriam Hernández Álvarez
C.I.: 1705009304

DECLARACIÓN DEL JEFE DE DEPARTAMENTO

Esta propuesta ha sido aprobada y avalada por el Consejo del Departamento de Informática y Ciencias de la Computación, en sesión del día 29 de agosto de 2017 mediante resolución No. 121.032.29-08-2017.

Las instalaciones, incluyendo personal, edificios, equipo y recursos financieros están a disposición del proponente y sus colaboradores de acuerdo con las especificaciones que se encuentran en esta propuesta.



Firma del Jefe del Departamento
Nombre: Myriam Peñafiel
C.I.: 1705828711

