

UNIVERSIDAD NACIONAL DE TRUJILLO
Facultad de Ciencias Físicas y Matemáticas
Escuela Profesional de Informática



**PREDICCIÓN DE LA RESPUESTA CORRECTA DE UNA
PREGUNTA DE OPCIÓN MÚLTIPLE MEDIANTE
TÉCNICAS DE APRENDIZAJE NO SUPERVISADO
ASOCIADO CON LA EXPERIENCIA.**

Nombre de autor: José Vicente Clavo Tafur

Nombre del Asesor: Dr. Jorge Luis Gutierrez Gutierrez

Trujillo - La Libertad

2018

PROYECTO DE INVESTIGACIÓN PARA TRABAJO DE GRADUACIÓN

ESCUELA PROFESIONAL DE INFORMÁTICA

1. GENERALIDADES

1.1. Título

Predicción de la respuesta correcta de una pregunta de opción múltiple mediante técnicas de aprendizaje no supervisado asociado con la experiencia.

1.2. Autor(es)

Tabla 1: Datos del alumno investigador

Código(s)	Nombres y Apellidos	Cargo en el proyecto	Email
xxxxxxxx-xx	José V. Clavo Tafur	Estudiante invest.	jclavotafur@gmail.com

1.3. Tipo de investigación

1.3.1. De acuerdo al fin que se persigue (Básica/Aplicada):

Básica: También denominada pura o fundamental, busca el progreso científico, incrementa los conocimientos teóricos, sin interesarse en las posibles aplicaciones o consecuencias prácticas; es mas formal y persigue las generalizaciones con vistas al desarrollo de una teoría basada en principios y leyes.

Aplicada: Tiene relación con la investigación básica, pues depende de los descubrimientos y avances de la investigación básica y se enriquece de ellos, pero se caracteriza por su interés en la aplicación, utilización y consecuencias prácticas de los conocimientos.

1.3.2. De acuerdo al alcance de la investigación:

Investigaciones cuantitativas, segun el nivel de asociación de las variables principales Lara Muñoz (2013):

- Explicativa o causal:** Consiste en la manipulación de una (o más) variable experimental no comprobada, en condiciones rigurosamente controladas, con el fin de describir de qué modo o por qué causa se produce una situación o acontecimiento particular. El experimento provocado por el investigador le permite introducir determinadas variables de estudio manipuladas por él, para controlar el aumento o disminución de esas variables y su efecto en las conductas observadas. Su método es cuantitativo y su fin es el descubrimiento de las causas.

- b) **Exploratorio:** Es un estudio inicial (examinar un tema poco estudiado) de un fenómeno, normalmente buscando un primer conocimiento del mismo; por ejemplo, cuáles son los aspectos y variables más significativos del fenómeno; y los resultados se consideran provisionales y la base para investigaciones posteriores.
- c) **Descriptivo:** Busca describir (frecuencias, porcentajes, medias, etc.) sin relacionarlas inferencialmente, las variables del fenómeno que se estudia. Se sitúa en el presente, pero no solamente se limita a la simple recolección y tabulación de datos, sino que hace la interpretación y el análisis imparcial de los mismo.
- d) **Correlacional o asociativa:** Busca determinar qué cambios de unas variables están asociados (correlacionados) con cambios en otras variables, normalmente sin establecer relaciones de causalidad, como sí es el caso en los análisis estructurales.

1.4. Área y línea de Investigación

1.4.1. Área de investigación :

Ejemplo: Algoritmos y complejidad.

1.4.2. Línea de Investigación:

Ejemplo: Estrategias algorítmicas.

1.4.3. Tema de investigación :

qwqwqqwq

1.5. Localidad e Institución donde se desarrollará el proyecto

1.5.1. Localidad (Dirección, Distrito, Provincia, Departamento) :

1.5.2. Institución (Universidad/Facultad/Departamento):

1.6. Duración del trabajo de graduación (Plan TG y desarrollo del TG)

Del 04/02/2019 AL 28/06/2019 (4. meses + 25. días)

1.7. Cronograma del trabajo de graduación

1.8. Recursos disponibles

1.8.1. Personal:

Personal técnico, administrativo y de servicios disponibles para el proyecto.

1.8.2. Materiales y Equipos:

Se debe especificar la calidad y cantidad de equipos, instrumentos y materiales disponibles para ejecutar el trabajo de investigación.

Tabla 2: Etapas y actividades para el trabajo de graduación

Etapas	Actividades/tareas	Fecha inicio	Fecha término	Hs. semanal
Preparación del plan TG.	Elaborar plan TG.	05/08/2016	05/05/2017	4
	Aprob. plan TG.	20/09/2016	05/05/2017	1
Recopilación de información.	Inv. bibliográfica.	05/05/2017	05/05/2017	30
	Instrumentos de medición.	05/05/2017	05/05/2017	25
Análisis de datos recolectados.	Procesamiento e interpretación de la información.	05/05/2017	05/05/2017	10
Resultados.	Desarrollo de TG según cronograma.	05/05/2017	05/05/2017	10
Redacción del informe TG.	Informe N1: Marco teórico y metodología.	05/05/2017	05/02/2017	50
	Informe N2: Propuesta de solución al problema formulado.	05/07/2017	05/05/2017	93
	Informe N3: Resultados y conclusiones preliminares.	05/08/2017	05/06/2017	20
Recopilación de información adicional para la tesis	Inv. bibliográfica adicional y su interpretación.	05/08/2016	05/05/2017	4
Pruebas y análisis de resultados	Preparación de resultados finales.	05/08/2016	05/05/2017	4
Redacción del informe de tesis	Informe de tesis de acuerdo con formato establecido.	05/08/2016	05/05/2017	4

Fuente: Elaboración propia.

1.8.3. Locales:

Señalar los ambientes donde se realizará la investigación: laboratorios, aulas, biblioteca, hemeroteca, etc. indicando su ubicación precisa.

1.9. Presupuesto

Costear en nuevos soles, los bienes, servicios e inversiones necesarios para llevar a cabo la investigación y que no estén disponibles. Presentar ordenados de acuerdo a la codificación del Clasificador de Gastos vigente. Considerar calidad, cantidad y precio.

1.10. Financiamiento

1.10.1. Con recursos universitarios:

Los recursos disponibles por subvención por investigación.

1.10.2. Con recursos externos:

Los que se perciban de fuente distinta a la UNT. Indicar la entidad aportante y los montos.

1.10.3. Autofinanciación:

Aporte del investigador (Puede ser en bienes o efectivo).

2. PLAN DE INVESTIGACIÓN

Es parte del proyecto de Investigación, debe ser lo suficientemente detallado para permitir comprender la naturaleza y los alcances de la investigación; así como, la rigidez del método seguido. Debe contener las citas bibliográficas (en norma APA) y estructurarse con subtítulos de acuerdo los ítems siguientes:

2.1. Realidad problemática

Para la realización de una investigación, el equipo investigador debe ubicarse en la realidad problemática dentro de un campo de interés y resaltar su importancia. Considerar estadísticas e indicadores numéricos que describan la situación problemática de la realidad.

2.2. Antecedentes

Para la realización de una investigación, el equipo investigador debe citar y comentar investigaciones recientemente realizadas que se relacionan con el problema o tema de investigación.

Ejemplo de antecedente: Cuando se expresa exactamente lo que dice el autor:

Ghiani et al. (2004) entienden que la logística trata de la planificación y control de los flujos de materiales e informaciones relacionadas en las organizaciones, tanto en los sectores público y privado. Además su misión es hacer la entrega de los productos correctos, en el local correcto y en la hora correcta, optimizando los costos operacionales totales del proceso. satisfaciendo un determinado conjunto de restricciones o condiciones.

Ejemplo de antecedente: Cuando se interpreta lo que dice el autor:

El desarrollo sustentable, (Figura 1), estará garantizado si se consideran tres aspectos fundamentales: económico, social y ambiental, donde la intersección de estos aspectos garantiza la calidad de vida en el espacio urbano y el equilibrio en las clases sociales en busca del bienestar (Tanguay et al., 2010).

Otros ejemplos de antecedentes:

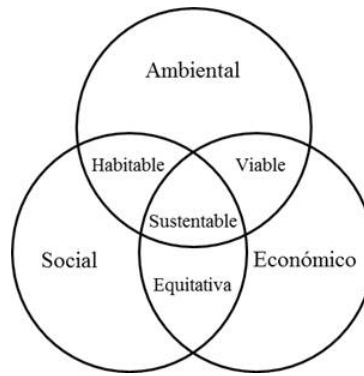


Figura 1: Aspectos claves para el desarrollo sustentable.

Fuente: Tanguay et al. (2010)

En los años 90 se presentaron definiciones generales las cuales vienen siendo mejoradas. Dekker et al. (2003) presenta una mejora en la definición de logística reversa como "el proceso de planificación, implementación y control de los flujos de materias-primas, en procesos de inventarios y bienes acabados, desde el punto de fabricación, distribución o uso, hacia el punto de recuperación o de eliminación".

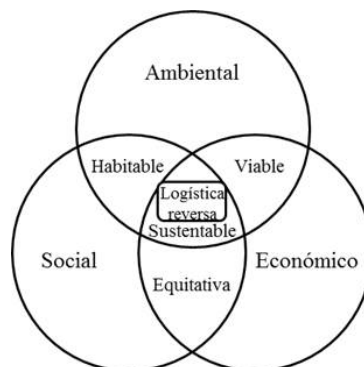


Figura 2: Logística reversa incluida en el desarrollo sustentable.

Fuente: Adaptación de Tanguay et al. (2010)

El problema de ruteo de vehículos (Ombuki et al., 2006; Yeun et al., 2008) y sus variantes han ganado mucho interés en la comunidad académica. La intención de estar más cerca a la realidad mediante el modelamiento matemático, hace que se hayan desarrollado nuevos modelos de optimización.

Según Sterle (2010) el primer nivel de la red comprende la distribución de la carga desde las plataformas hasta las unidades satélites, utilizando vehículos de carga de mayor tamaño (g). El segundo nivel, consiste en montar rutas desde las unidades satélites hasta los clientes, usando para este caso vehículos de menor tamaño (v). El modelo de localización y ruteo de vehículos para la distribución de carga propuesto por el autor, además de hacer la conexión de los dos niveles y estudiar su inter relación y dependencia, el modelo busca determinar la cantidad necesaria de plataforma y de unidades satélites considerando el tamaño y dimensionamiento de la

flota para el ruteo en dos niveles.

Tabla 3: Resultados computacionales obtenidos en el modelo de Sterle (2010)

Escenarios	Demanda cliente (ton.)	Tiempo (min.)	Costo (\$)
1	P1:1; P2:2; P3:2; P4:2; P5:1	0.12	667.42
2	P1:1; P2:2; P3:2; P4:2; P5:1; P:4; P7:3	56.54	1744.35
3	P1: 1; P2:2; P3:2; P4:2; P5:1; P6: 4; P7:3; P8:2; P9:2	287.70	1750.72
4	P1:1; P2:2; P3: 2; P4:2; P5:1; P6:4; P7:3; P8:2; P9:2; P10:1	1848.57	1773.46

Fuente: Resultados obtenidos con CPLEX.

2.3. Justificación

La mayoría de las investigaciones se efectúan con un propósito definido. Tal propósito debe ser lo suficientemente fuerte para que justifique su realización. Lara Muñoz (2013)

- (a) Razones o motivos e importancia del tema a ser investigado.
- (b) Sustentar la pertinencia de la pregunta o problema que se abordará en la investigación.
- (c) Considerar los resultados esperados e impactos previstos.

2.4. Problema

Plantear el problema no es otra cosa mas que afinar y estructurar formalmente la idea de investigación. El planteamiento del problema puede ser sencillo o complejo dependiendo de la familiarización del investigador en el tema a tratar.

Los criterios para formular un problema de investigación son:

- a) El problema debe ser formulado claramente y sin ambigüedad como pregunta: ¿Qué efecto?; ¿en qué condiciones ...?; ¿cuál es la probabilidad de ...?; ¿cómo se relaciona ... con ...?; ¿cómo ...?.
- b) El planteamiento debe implicar la posibilidad de realizar una prueba empírica o una recolección de datos, es decir, la factibilidad de observarse en la realidad o en un entorno.

Por lo tanto, los elementos para plantear un problema de investigación son tres y estan relacionados entre si: Los objetivos que persigue la investigación; las preguntas de investigación y la justificación del estudio. Lara Muñoz (2013)

Ejemplo:

¿Cómo viabilizar una red logística reversa en regiones urbanas minimizando los costos logísticos de ruteo y transporte de los RSU hasta su disposición final?

2.5. Hipótesis

Preferentemente para investigaciones explicativas debe ser una respuesta a priori y tentativa guardando coherencia con el problema científico, se formula como una proposición afirmativa, con un lenguaje claro y específico. Las hipótesis se obtienen por deducción lógica y está sustentada en los conocimientos científicos.

Criterios para formular hipótesis: Lara Muñoz (2013)

- a) Toda hipótesis de investigación debe ser verificable estadísticamente. Puede ser difícil o imposible de verificar porque no existe un conocimiento sobre el cual se pueda formular una hipótesis, o bien, porque una o más variables no son medibles.
- b) Toda hipótesis debe indicar la relación entre variables, lo que implica que las variables deben ser medibles.
- c) Toda hipótesis debe tener sus límites. Pueden escogerse hipótesis que sean sencillas de validar, y sin embargo, altamente significativas.
- d) El investigador debe tener una razón específica para considerar una hipótesis, ya sea teórica o por alguna evidencia concreta.

2.6. Variables

Hipótesis: Existe un mayor número de plantas comestibles en climas cálidos que en climas fríos.

Los elementos que se están relacionando son:

- (1) plantas comestibles
- (2) climas cálidos
- (3) climas fríos

Estos tres elementos son variables.

2.6.1. Variable independiente

Los climas.

2.6.2. Variables dependiente

Plantas comestibles.

2.7. Objetivos

2.7.1. Objetivos generales

- a) Desarrollar un algoritmo de aprendizaje no supervisado para predecir la respuesta correcta de una pregunta de opción múltiple.

2.7.2. Objetivos específicos

- a) Codificar el algoritmo de aprendizaje no supervisado para trabajar con clusters.
- b) Desarrollar del prototipo de software y aplicarla a la predicción de la respuesta correcta.
- c) Realizar las pruebas y documentar resultados.

2.8. Método de trabajo

Para llegar a los objetivos propuestos, el desarrollo de la investigación comprendió las siguientes etapas de trabajo a saber:

- a) Formulación del problema principal de la investigación, justificando su importancia.
- b) Búsqueda del material bibliográfico de los diferentes temas necesarios para la elaboración de la investigación, tales como aprendizaje automático no supervisado, probabilidades, estadísticas, desarrollo web, entre otros.
- c) Estudio y análisis de los algoritmos de aprendizaje automático no supervisado.
- d) Estudio de modelos probabilísticos y estadísticos que contribuyan con el tema de predicciones.
- e) Desarrollo del algoritmo de aprendizaje automático no supervisado.
- f) Desarrollo del prototipo de software web.
- g) Elección de los datos de prueba.
- h) Testear, validar y documentar los resultados de la investigación

2.9. Referencias

Presentar bibliografía conforme a las normas técnicas internacionales reconocidas: Un solo estilo American Psychological Association - APA. Seguir exactamente la forma de las referencias dadas como ejemplos. No usar numeros entre corchetes. Ver los ejemplos de las referencias dadas abajo.

Se exige como mínimo 10 referencias entre libros y/o artículos referentes al tema a investigar. Las referencias sustentan la investigación del TG.

Referencias

- Dekker, R., Inderfurth, K., van Wassenhove, L., and Fleischmann, M. (2003). Quantitative approaches for reverse logistics. *Springer-Verlag, Berlin. Forthcoming.*, 24(22):40–60.
- Ghiani, G., Laporte, G., and Musmanno, R. (2004). *Introduction to logistics systems planning and control*. Jhon Wiley & Sons.

- Korte, B. and Vygen, J. (2008). *Combinatorial optimization Theory and Algorithms, Fourth Edition*. Springer Heidelberg Dordrecht., London New York.
- Lara Muñoz, E. M. (2013). *Fundamentos de investigación: Un enfoque por competencias, 2da. edición*. Alfaomega, México.
- Ombuki, B., Ross, B. J., and Hanshar, F. (2006). Multi-objective genetic algorithms for vehicle routing problem with time windows. *Applied Intelligence, Springer Science.*, 322(12):91–99.
- Rojas Soriano, R. (2001). *Guía para realizar investigaciones sociales, 26a. edición*. Plaza Valdez, México.
- Sterle, C. (2010). *Location-Routing models and methods for freight distribution and infomobility in city logistics*. Thèse de doctorat. Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport., CIRRELT.
- Tanguay, G., Rajaonson, J., Lefevre, J., and Lanoie, P. (2010). Measuring the sustainability of cities: An analysis of the use of local indicators. *Ecological Indicators, Elsevier.*, 26(22):100–115.
- Yeun, L., Ismail, W., Omar, K., and Zirour, M. (2008). Vehicle routing problem: Models and solutions. *Journal of Quality measurement and analysis.*, 22(13):111–120.

Edgar M. Peche Perlado
Alumno

Manuel E. Pérez Yon
Alumno

Alan M. Turing Godel
Asesor