



ANEXO 1

FORMATO PARA LA PRESENTACIÓN DE PROYECTOS DE INVESTIGACIÓN CON EL FINANCIAMIENTO DEL FEDU

1. Título del proyecto

Detección de correo no deseado utilizando aprendizaje de máquina en Octave

2. Área de Investigación

Área de investigación	Línea de Investigación	Disciplina OCDE
Ciencia de datos	Machine learning	

3. Duración del proyecto (meses)

12 meses

4. Tipo de proyecto

Individual	<input checked="" type="radio"/>
Multidisciplinario	<input type="radio"/>
Director de tesis pregrado	<input type="radio"/>

4. Datos de los integrantes del proyecto

Apellidos y Nombres	1. VILCA HUAYTA OLIVER AMADEO
Escuela Profesional	1. INGENIERÍA DE SISTEMAS
Celular	1. 966107864
Correo Electrónico	1. ovilca@unap.edu.pe

- I. **Título** (El proyecto de tesis debe llevar un título que exprese en forma sintética su contenido, haciendo referencia en lo posible, al resultado final que se pretende lograr. Máx. palabras 25)

CURVA EPIDÉMICA Y CANAL ENDÉMICO DE LA ENFERMEDAD COVID-19 PERÚ 2021

- II. **Resumen del Proyecto de Tesis** (Debe ser suficientemente informativo, presentando -igual que un trabajo científico- una descripción de los principales puntos que se abordarán, objetivos, metodología y resultados que se esperan)

A pesar del desarrollo actual en detección de correo no deseado, el adversario sigue arreglándoselas para pasar desapercibido y lograr sus objetivos, por lo que es necesario continuar con las mejoras. El proyecto de investigación tiene el objetivo de construir un detector de correo no deseado utilizando el aprendizaje de máquina y mediante el entrenamiento con una base de datos de mensajes de correos electrónico. La base de datos que se utilizará será una de las actuales que este a disposición. Se empleará el aprendizaje de máquina, específicamente una adaptación de la regresión logística y otros que sean adecuados. Finalizada la investigación se



logrará un detector de correo no deseado con sus respectivos indicadores de exactitud.

III. Palabras claves (Keywords) (Colocadas en orden de importancia. Máx. palabras: cinco)

correo no deseado, aprendizaje de máquina, Octave

IV. Justificación del proyecto (Describa el problema y su relevancia como objeto de investigación. Es importante una clara definición y delimitación del problema que abordará la investigación, ya que temas cuya definición es difusa o amplísima son difíciles de evaluar y desarrollar)

Tan pronto como los mensajes de correo electrónico crecen, es necesario implementar y/o mejorar los sistemas de detección de correos no deseados (spam). Entre los primeros beneficios que se tienen es evitar las estafas y otros que suelen aparecer en los noticieros.

V. Antecedentes del proyecto (Incluya el estado actual del conocimiento en el ámbito nacional e internacional. La revisión bibliográfica debe incluir en lo posible artículos científicos actuales, para evidenciar el conocimiento existente y el aporte de la Tesis propuesta. Esto es importante para el futuro artículo que resultará como producto de este trabajo)

Actualmente, se están realizando muchas investigaciones a fin de implementar sistemas de detección de correo no deseado. De hecho, lo tienen los proveedores de servicio de correo electrónico como Gmail, Outlook, Yahoo! y los servidores de correo electrónico que prestan servicios de mensajes. En relación a investigaciones cada año se incrementan publicaciones en este ámbito.

VI. Hipótesis del trabajo (Es el aporte proyectado de la investigación en la solución del problema)

VII. Objetivo general

Desarrollar un detector de correo no deseado.

VIII. Objetivos específicos

Implementar un sistema preprocesamiento.
Entrenar con la base de datos.
Evaluar el desempeño de sistema de detección de correo no deseado

IX. Metodología de investigación (Describir el(los) método(s) científico(s) que se empleará(n) para alcanzar los objetivos específicos, en forma coherente a la hipótesis de la investigación. Sustentar, con base bibliográfica, la pertinencia del(los) método(s) en términos de la representatividad de la muestra y de los resultados que se esperan alcanzar. Incluir los análisis estadísticos a utilizar)

MATERIAL
Bases de datos



FUENTE DE INFORMACIÓN

Reportes de publicaciones en páginas Web.

MÉTODO

Regresión logística en principio y otros

X. Referencias (Listar las citas bibliográficas con el estilo adecuado a su especialidad)

XI. Uso de los resultados y contribuciones del proyecto (Señalar el posible uso de los resultados y la contribución de los mismos)

El detector de spam contribuirá a detectar correo no deseado en diferentes sistemas que lo requieran.

XII. Impactos esperados

i. Impactos en Ciencia y Tecnología

El desarrollo de un sistema de detección de correo no deseado

ii. Impactos económicos

Costos que se evitarían al disponerse de un sistema mejorado de detección de correo no deseado

iii. Impactos sociales

Varios que se verán.

iv. Impactos ambientales

Indirectamente

XIII. Recursos necesarios (Infraestructura, equipos y principales tecnologías en uso relacionadas con la temática del proyecto, señale medios y recursos para realizar el proyecto)

HUMANOS: Los autores.
MATERIALES: Varios.

XIV. Localización del proyecto (indicar donde se llevará a cabo el proyecto)

El proyecto se realizará en los ambientes de la Universidad Nacional del Altiplano Puno.

XV. Cronograma de actividades

Actividad	Trimestres											
	E N E	F E B	M A R	A B R	M A Y	J U N	J U L	A G O	S E P	O C T	N O V	D I C

