**UNIVERSIDAD PRIVADA DE TACNA**

**FACULTAD DE INGENIERÍA**

**Escuela Profesional de Ingeniería de Sistemas**

****

**“Sistema web con integración de Machine Learning para la detección anticipada de keyloggers en instituciones educativas - 2025”**

Curso: ***Construcción de Software I***

Docente: ***ING. ALBERTO JONATAN FLOR RODRIGUEZ***

Integrantes:

**Arce Bracamonte, Sebastian Rodrigo (2019092986)**

**Tacna – Perú**

***2025***

| CONTROL DE VERSIONES | | | | | |
| --- | --- | --- | --- | --- | --- |
| Versión | Hecha por | Revisada por | Aprobada por | Fecha | Motivo |
| 1.0 | SAB | SAB | SAB | 11/08/2025 | Versión Original |

**“Sistema web con integración de Machine Learning para la detección anticipada de keyloggers en instituciones educativas - 2025”**

**Propuesta de Proyecto**

**Versión *1.0***

| CONTROL DE VERSIONES | | | | | |
| --- | --- | --- | --- | --- | --- |
| Versión | Hecha por | Revisada por | Aprobada por | Fecha | Motivo |
| 1.0 | SAB | SAB | SAB | 11/08/2025 | Versión Original |

**Índice**

[**1. Planteamiento del problema 4**](#_heading=h.j5ns9hezvham)

[**2. Justificación del proyecto 5**](#_heading=h.rsro9yhlz1ch)

[2.1. Vulnerabilidad Creciente y Amenazas Activas 5](#_heading=h.dw3lpifctugu)

[2.2. Protección de Datos Sensibles y Cumplimiento Legal 5](#_heading=h.2f44b8e727cr)

[2.3. Insuficiencia de las Soluciones Tradicionales 5](#_heading=h.stfqt0iwwf7p)

[2.4. Impacto Económico Directo y Retorno de Inversión (ROI) 5](#_heading=h.rxx1wpt4xbzh)

[2.5. Ventaja Competitiva y Posicionamiento Institucional 6](#_heading=h.bktoxrhprmbz)

[2.6. Desarrollo de Capacidades Locales y Concienciación 6](#_heading=h.ivrmavq2ds37)

[2.7. Alineamiento con Estrategias Nacionales de Digitalización 6](#_heading=h.dtovljrbh68r)

[**3. Objetivo del proyecto 6**](#_heading=h.c3zqeyqkn3d2)

[3.1. Objetivo General 6](#_heading=h.bya3qh14fb7r)

[3.2. Objetivos Específicos 6](#_heading=h.kynewcbfvsc4)

[**4. Beneficios 7**](#_heading=h.say21au3ipvz)

[4.1. Beneficios Tangibles (Cuantificables) 7](#_heading=h.stxz9cnq54uz)

[4.2. Beneficios Intangibles (Estratégicos) 8](#_heading=h.pp7ydelczj1n)

[4.2.1. Fortalecimiento de la Seguridad 8](#_heading=h.jihlr7oikv5y)

[4.2.2. Ventaja Competitiva Institucional 8](#_heading=h.600pygcs67it)

[4.2.3. Cumplimiento Normativo 8](#_heading=h.kxukmop3hraj)

[4.3. Beneficios Operativos 8](#_heading=h.o8sfcmqd24rq)

[4.3.1. Eficiencia Operativa 8](#_heading=h.f91kxpl3ccer)

[4.3.2. Capacitación y Conocimiento 8](#_heading=h.qep9mlqheuc7)

[4.4. Beneficios Sociales y Educativos 9](#_heading=h.mnibxd4em8jd)

[4.4.1. Impacto en la Comunidad Educativa 9](#_heading=h.xp8hhscfcfco)

[4.4.2. Desarrollo de Capacidades Locales 9](#_heading=h.u2413jkpxdnm)

[4.5. Beneficios Estratégicos Nacionales 9](#_heading=h.ro8tejj6ypnx)

[4.5.1. Alineamiento con Políticas Públicas 9](#_heading=h.gjg4wlh1oub6)

[4.5.2. Sostenibilidad del Proyecto 9](#_heading=h.miudmdi4pn4m)

[4.6. Resumen de Valor Total 9](#_heading=h.urz357svas09)

[**5. Alcance 10**](#_heading=h.hfy4p087vr0a)

[1. Desarrollo del Aplicativo Web 10](#_heading=h.obzkz2ctffuc)

[2. Implementación de Machine Learning 10](#_heading=h.ynpmbosw2vnj)

[3. Capacitación y Sensibilización 10](#_heading=h.47s0eyvw7w9w)

[4. Monitoreo y Evaluación 10](#_heading=h.ufcz9ayxnrm0)

[5. Alcance Geográfico 11](#_heading=h.jy4lut955xti)

[6. Resultados Esperados 11](#_heading=h.yj2j6an1bimr)

[**6. Requerimientos del sistema 11**](#_heading=h.otafblu3ppyq)

[**7. Restricciones 11**](#_heading=h.4bn95wyg63ix)

[**8. Supuestos 11**](#_heading=h.sykspb4g80kz)

[**9. Resultados Esperados 11**](#_heading=h.ph2xhxqfo641)

[**10. Metodología de implementación 11**](#_heading=h.pbvdp5pxj3gz)

[**11. Actores claves 11**](#_heading=h.q4qp5ib2qdjm)

[**12. Papel y responsabilidad personal 12**](#_heading=h.s2mp73qrawd3)

[**13. Plan de monitoreo y evaluación 12**](#_heading=h.2h77q0ciys7l)

[**14. Cronograma del proyecto 13**](#_heading=h.9t66chdtsl7o)

[**15. Hitos de entregable 13**](#_heading=h.5zz9g2sxwu)

[**RESUMEN EJECUTIVO 15**](#_heading=h.6e3amk7a1bv)

# Planteamiento del problema

En la era digital actual, las instituciones educativas de Tacna, al igual que en el resto del Perú, han acelerado su transformación digital, integrando plataformas en línea para la gestión académica, administrativa y financiera. Si bien esta conectividad ofrece enormes beneficios, también expande la superficie de ataque para ciberamenazas sofisticadas. Entre estas, los keyloggers se han posicionado como una de las amenazas más críticas y silenciosas. Este malware, diseñado para capturar las pulsaciones del teclado, compromete información sensible como credenciales de acceso, datos personales de estudiantes y docentes, e incluso detalles financieros asociados a pagos de matrículas o becas.

La problemática se agrava por la dependencia de soluciones de seguridad tradicionales (como antivirus basados en firmas) que resultan insuficientes frente a keyloggers polimórficos o de día cero (zero-day), capaces de evadir detección. Según un reporte de la Divindat (2024), el 60% de las instituciones educativas en el sur del Perú carecen de protocolos avanzados de ciberseguridad, relying principalmente en medidas reactivas que solo actúan ante amenazas ya conocidas, dejando un vacío de protección contra ataques noveles.

El factor humano amplifica este riesgo. La falta de capacitación en higiene digital entre estudiantes, docentes y personal administrativo conduce a prácticas de riesgo, como hacer clic en enlaces maliciosos o descargar software no verificado, que son vectores comunes de infección para keyloggers.

Las consecuencias de un incidente exitoso son severas y multifacéticas:

* Económicas: Pérdidas directas por fraudes financieros. instituciones en Tacna reportaron pérdidas por suplantación de identidad que superaron los S/ 120,000 en los últimos dos años.
* Legales: Potencial incumplimiento de la Ley N° 29733 (Ley de Protección de Datos Personales), que puede acarrear sanciones económicas significativas y daños a la reputación institucional.
* Operativas: Interrupción de servicios educativos, pérdida de información académica crítica y saturación de los equipos de TI con labores de remediation.
* Reputacional: La pérdida de confianza por parte de estudiantes, familias y la comunidad en general puede traducirse en una disminución de la matrícula y un debilitamiento de la imagen institucional.

En este contexto, se hace evidente la necesidad crítica de desarrollar e implementar un sistema de detección anticipada proactivo. Dicho sistema debe superar las limitaciones de los métodos convencionales mediante el uso de técnicas de Machine Learning (ML) que permitan analizar patrones de comportamiento del sistema y del usuario en tiempo real, identificando anomalías sutiles que delaten la presencia de un keylogger, incluso si es una variante nunca antes vista.

# Justificación del proyecto

La implementación de un sistema avanzado de detección de keyloggers mediante Machine Learning se justifica en el contexto educativo actual por las siguientes razones críticas, que abarcan aspectos normativos, económicos, operativos y estratégicos:

#### 2.1. Vulnerabilidad Creciente y Amenazas Activas

El sector educativo se ha convertido en un blanco frecuente de ciberataques. Según el CERT-PE (Centro de Operaciones de Seguridad del Perú), en 2023 se reportó un incremento del 40% en incidentes de seguridad contra universidades e institutos superiores a nivel nacional, siendo el robo de credenciales mediante keyloggers una de las técnicas más recurrentes. Esta tendencia confirma la urgencia de implementar soluciones proactivas que anticipen y neutralicen amenazas antes de que materialicen su impacto.

#### 2.2. Protección de Datos Sensibles y Cumplimiento Legal

Las instituciones educativas son custodias de masivos volúmenes de datos sensibles (registros académicos, información financiera, datos personales de menores de edad). Su protección no es solo una prioridad operativa, sino un imperativo legal bajo la Ley N° 29733 (Ley de Protección de Datos Personales) y su Reglamento (Decreto Supremo N° 003-2013-JUS). La no implementación de medidas de seguridad adecuadas puede acarrear sanciones de INDECOPI de hasta 100 UIT (S/ 495,000), además de demandas por daños y perjuicios. Este proyecto constituye una inversión en cumplimiento normativo y mitigación de riesgo legal.

#### 2.3. Insuficiencia de las Soluciones Tradicionales

Las defensas perimetrales tradicionales (firewalls, antivirus basados en firmas) han demostrado ser incapaces de detectar keyloggers avanzados. Estos malware suelen utilizar técnicas de ofuscación, cifrado y polimorfismo que eluden las listas de firmas conocidas. La propuesta de utilizar Machine Learning representa un cambio de paradigma: de una detección reactiva a una detección basada en comportamientos y anomalías, permitiendo identificar amenazas noveles o variantes de keyloggers nunca antes vistas con una tasa de efectividad superior al 95%.

#### 2.4. Impacto Económico Directo y Retorno de Inversión (ROI)

Los costos asociados a un incidente de seguridad son sustanciales. Incluyen:

* Costos de remediación: Investigación, contención, recuperación de datos y restablecimiento de servicios.
* Multas y sanciones: Por incumplimiento de la normativa de protección de datos.
* Daño reputacional: Pérdida de matrículas y deserción estudiantil.

La implementación de este sistema, con una inversión estimada de S/ 101,860, prevendría pérdidas potenciales calculadas en S/ 120,000 anuales para una institución mediana, ofreciendo un Retorno de Inversión (ROI) claro y un período de recuperación de apenas 2.1 años.

#### 2.5. Ventaja Competitiva y Posicionamiento Institucional

En un mercado educativo cada vez más competitivo, la ciberseguridad se ha convertido en un diferenciador estratégico. Las instituciones que demuestren un compromiso tangible con la protección de sus estudiantes y colaboradores mediante tecnologías de vanguardia mejorarán su percepción pública, atraerán y retendrán talento, y se posicionarán como líderes innovadores en el sector educativo de Tacna.

#### 2.6. Desarrollo de Capacidades Locales y Concienciación

El proyecto trasciende la instalación de un software. Incluye un componente fundamental de capacitación que busca:

* Fortalecer las capacidades del capital humano local en técnicas modernas de ciberseguridad e inteligencia artificial.
* Promover una cultura de seguridad proactiva entre estudiantes, docentes y personal administrativo.
* Crear un modelo replicable que pueda ser adoptado por otras instituciones de la región, contribuyendo a la resiliencia digital del ecosistema educativo surperuano.

#### 2.7. Alineamiento con Estrategias Nacionales de Digitalización

Esta iniciativa se alinea directamente con los objetivos del Plan Nacional de Transformación Digital del Perú y las prioridades del Plan Nacional de Ciberseguridad, que destacan la necesidad de proteger la infraestructura digital crítica del país, dentro de la cual el sector educativo es un pilar fundamental.

# Objetivo del proyecto

### 3.1. Objetivo General

Desarrollar e implementar un sistema web basado en técnicas de Machine Learning para la detección proactiva y neutralización de keyloggers en las redes informáticas de instituciones educativas de Tacna, que fortalezca la seguridad de la información, proteja datos sensibles y sirva como modelo de referencia en ciberseguridad para el sector educativo peruano.

### 3.2. Objetivos Específicos

* + 1. Recolectar y procesar datos de comportamiento de red y procesos para crear un dataset robusto y representativo que permita el entrenamiento de modelos de Machine Learning capaces de identificar patrones anómalos asociados a keyloggers.
    2. Diseñar, entrenar y validar algoritmos de Machine Learning (como Redes Neuronales Recurrentes - RNN o Random Forests) que logren una tasa de detección superior al 95% con menos del 3% de falsos positivos en la identificación de actividad maliciosa de keylogging.
    3. Desarrollar una plataforma web intuitiva y centralizada que integre los modelos de ML, permitiendo el monitoreo en tiempo real, la visualización de alertas y la gestión de incidentes para el personal técnico de las instituciones.
    4. Validar el sistema en entornos reales controlados mediante la ejecución de simulaciones de ataques con keyloggers de diferentes niveles de sofisticación, demostrando una efectividad de neutralización del 90% en las amenazas identificadas.
    5. Capacitar al personal técnico y administrativo de al menos 3 instituciones educativas piloto en Tacna sobre el uso, mantenimiento y interpretación de alertas del sistema, garantizando su operatividad autónoma y la correcta respuesta ante incidentes.
    6. Documentar el proceso de desarrollo, los costos asociados y el análisis de riesgos del proyecto, generando un modelo escalable y replicable que pueda ser adoptado por otras instituciones educativas a nivel nacional.
    7. Realizar un análisis de costo-beneficio que evidencie el retorno de la inversión (ROI) y la reducción de costos operativos asociados a la remediation de incidentes de seguridad, estableciendo los fundamentos para la sostenibilidad del proyecto a largo plazo.

# Beneficios

### 4.1. Beneficios Tangibles (Cuantificables)

| Beneficio | Impacto Esperado | Valor Estimado (S/) |
| --- | --- | --- |
| Reducción de costos por incidentes de seguridad | Disminución del 80% en pérdidas por fraudes y remediation | S/ 25,000 anuales |
| Ahorro en licencias de seguridad | Sustitución de herramientas comerciales por solución propia | S/ 12,000 anuales |
| Optimización de tiempo del personal de TI | Reducción de 70% en horas dedicadas a monitoreo manual | S/ 18,000 anuales |
| Prevención de multas por protección de datos | Cumplimiento de Ley N° 29733 evitando sanciones de INDECOPI | Hasta S/ 495,000 |
| Generación de ingresos por consultoría | Venta de servicios a otras instituciones educativas | S/ 15,000 anuales |

### 4.2. Beneficios Intangibles (Estratégicos)

#### 4.2.1. Fortalecimiento de la Seguridad

* Detección proactiva de keyloggers antes que comprometan datos sensibles
* Protección de información crítica: credenciales, datos académicos, información financiera
* Capacidad de adaptación continua frente nuevas variantes de malware
* Reducción de ventanas de vulnerabilidad de días a segundos

#### 4.2.2. Ventaja Competitiva Institucional

* Posicionamiento como líder en innovación tecnológica educativa
* Diferenciación en el mercado educativo regional y nacional
* Atracción de estudiantes que valoran la seguridad digital
* Retención de talento docente y administrativo

#### 4.2.3. Cumplimiento Normativo

* Aseguramiento de cumplimiento con Ley N° 29733 de protección de datos
* Preparación para regulaciones futuras en ciberseguridad
* Documentación y auditoría automática de incidentes de seguridad

### 4.3. Beneficios Operativos

#### 4.3.1. Eficiencia Operativa

* Centralización de la monitorización de seguridad en plataforma única
* Automatización de detección y respuesta a incidentes
* Reducción de falsos positivos del 30% al 3%
* Tiempo real de respuesta menor a 5 segundos

#### 4.3.2. Capacitación y Conocimiento

* Desarrollo de competencias en ML y ciberseguridad del personal local
* Transferencia de conocimiento a estudiantes mediante pasantías
* Creación de manuales y protocolos estandarizados
* Formación de multiplicadores en seguridad informática

### 4.4. Beneficios Sociales y Educativos

#### 4.4.1. Impacto en la Comunidad Educativa

* Protección de datos de menores de edad y población vulnerable
* Cultura de ciberseguridad extendida a estudiantes, docentes y familias
* Confianza digital en el uso de plataformas educativas
* Prevención de acoso digital y robo de identidad

#### 4.4.2. Desarrollo de Capacidades Locales

* Fortalecimiento del ecosistema tecnológico tacneño
* Generación de empleo especializado en ciberseguridad
* Modelo replicable para otras regiones del país
* Atracción de inversión en tecnología a la región

### 4.5. Beneficios Estratégicos Nacionales

#### 4.5.1. Alineamiento con Políticas Públicas

* Contribución al Plan Nacional de Ciberseguridad 2023-2026
* Apoyo a la política de transformación digital educativa
* Fortalecimiento de la soberanía digital peruana
* Preparación para ciberamenazas evolucionadas

#### 4.5.2. Sostenibilidad del Proyecto

* Modelo autosostenible mediante ahorros generados
* Escalabilidad a bajo costo para nuevas instituciones
* Actualización continua con mínima inversión adicional
* Comunidad de desarrollo open source alrededor del proyecto

### 4.6. Resumen de Valor Total

| **Dimensión** | **Beneficio Principal** | **Impacto** |
| --- | --- | --- |
| **Económica** | ROI del 28% anual | Retorno en 2.1 años |
| **Operativa** | Detección en <5 segundos | 95% de efectividad |
| **Estratégica** | Liderazgo en innovación | Posicionamiento único |
| **Social** | Protección comunidad educativa | 3,000+ usuarios protegidos |
| **Legal** | Cumplimiento normativo | Evita multas hasta 100 UIT |

Valor total estimado en 5 años: S/ 350,000 (beneficios tangibles + intangibles)

# Alcance

### 1. Desarrollo del Aplicativo Web

* Creación de un Sistema Accesible: Se desarrollará un sistema web que permita la gestión de la seguridad informática, accesible para el personal de TI, administrativos y usuarios finales. Esto incluirá una interfaz fácil de usar que facilite la interacción entre los diferentes roles.
* Integración de Módulos Específicos: El sistema incorporará módulos que implementen las funcionalidades clave de detección de amenazas y gestión de incidentes, adaptándose a las necesidades específicas de las instituciones educativas.

### 2. Implementación de Machine Learning

* Uso de Algoritmos Avanzados: El proyecto utilizará algoritmos de Machine Learning, como Random Forest y otros modelos supervisados, para analizar datos de tráfico y detectar patrones que puedan indicar amenazas informáticas.
* Tácticas Ofensivas Controladas: Se desarrollarán tácticas ofensivas que simulen ataques cibernéticos para evaluar la capacidad de respuesta del sistema ante diversas modalidades de intrusión, sin comprometer la seguridad del entorno educativo.

### 3. Capacitación y Sensibilización

* Entrenamiento del Personal: Se llevarán a cabo sesiones de capacitación para el personal de TI y seguridad en el uso efectivo del sistema, garantizando que cuenten con las habilidades necesarias para gestionar y operar la plataforma.
* Talleres de Ciberseguridad: Se organizarán talleres y campañas de sensibilización sobre ciberseguridad para toda la comunidad educativa, fomentando una cultura de prevención y preparación ante ciberamenazas.

### 4. Monitoreo y Evaluación

* Establecimiento de un Plan de Monitoreo: Se implementará un plan de monitoreo continuo que permita la recopilación de datos sobre la eficacia del sistema en la detección y respuesta a amenazas, así como la generación de informes periódicos sobre incidentes de seguridad.
* Evaluación del Sistema: Se realizará una evaluación continua del rendimiento del sistema y su efectividad en la detección de amenazas, con el objetivo de realizar ajustes necesarios para mejorar su funcionamiento.

### 5. Alcance Geográfico

* Instituciones Educativas en Tacna: El proyecto se llevará a cabo en diversas instituciones educativas de la ciudad de Tacna, con la posibilidad de expandirse a otras instituciones de la región en el futuro si se demuestra su efectividad.
* Limitaciones del Proyecto: Este proyecto se enfocará exclusivamente en la detección y neutralización de amenazas informáticas dentro de las redes educativas. No se incluirán otras áreas como la protección física de las instalaciones o la gestión de incidentes fuera del ámbito digital.

### 6. Resultados Esperados

* + Ambiente Educativo Más Seguro: Se espera que el sistema contribuya a la creación de un entorno educativo más seguro y confiable, donde estudiantes y personal se sientan protegidos ante amenazas cibernéticas.
  + Optimización de Recursos: La automatización de procesos de detección y respuesta permitirá una mejor asignación de recursos, facilitando que el personal se concentre en tareas de mayor valor estratégico.

# Requerimientos del sistema

### 1. Requisitos de Software

#### 1.1. Mínimos

* Sistema Operativo:
  + Windows 10 (64-bit)
  + Linux Ubuntu 20.04+
  + macOS 12+ (opcional)
* Entorno Python:
  + Python 3.8+
  + pip 20.0+
  + venv/virtualenv
* Paquetes Esenciales:
  + pandas >= 1.3.0
  + numpy >= 1.21.0
  + scikit-learn >= 1.0.0
  + onnxruntime >= 1.22.2 (CPU inference)
  + pymongo >= 4.0 (si usa MongoDB)
  + mysql-connector-python >= 8.0 (si usa MySQL)
* Herramientas Adicionales:
  + .NET Runtime 6.0+ (para componentes C#)
  + Visual Studio 2019/2022 (workload C++) o CMake >= 3.20
  + git >= 2.30

#### 1.2. Recomendados

* Python 3.10+
* onnxruntime 1.22.2 (GPU version con CUDA 11.6+)
* .NET SDK 7.0+
* MySQL Server 8.0+ o MongoDB 6.0+
* PowerShell 7.0+ (Windows) o bash 5.0+ (Linux/Mac)

### 2. Requisitos de Hardware

#### 2.1. Configuración Mínima

* CPU: 2 núcleos (x86\_64 architecture)
* RAM: 8 GB DDR4
* Almacenamiento: 5 GB SSD/HDD
* Red: Ethernet 100Mbps

#### 2.2. Configuración Recomendada

* CPU: 4+ núcleos (Intel i7/AMD Ryzen 7 o superior)
* RAM: 16 GB DDR4
* Almacenamiento: 50 GB SSD NVMe
* GPU: NVIDIA RTX 3060+ (12GB VRAM) con CUDA support
* Red: Gigabit Ethernet

### 3. Requisitos Funcionales

#### 3.1. Procesamiento de Datos

* Ingesta: Soporte para formatos CSV/XLSX desde directorio DATOS
* Preprocesamiento:
  + Normalización automática de datos
  + Manejo de valores faltantes
  + Selección de características numéricas
* Validación: Verificación de integridad de datos de entrada

#### 3.2. Machine Learning

* Entrenamiento:
  + Implementación de Random Forest y otros algoritmos
  + Guardado de artefactos (.pkl, .json features)
* Conversión: Exportación a formato ONNX (.onnx)
* Inferencia:
  + Carga de modelos pre-entrenados
  + Predicción en nuevos datos
  + Soporte para Python, C++ y C#

#### 3.3. Gestión de Modelos

* Persistencia:
  + Almacenamiento local de modelos
  + Integración opcional con MySQL/MongoDB
* Versioning:
  + Control de versiones de modelos
  + Metadata de entrenamiento

#### 3.4. Operaciones

* Batch Processing:
  + Procesamiento por lotes de múltiples archivos
  + Generación de reportes automáticos
* Logging:
  + Registro de ejecución y errores
  + Auditoría de operaciones
* CLI: Interfaces de línea de comandos para:
  + Entrenamiento (train\_from\_datos.py)
  + Predicción (predecir\_keylogger.py)
  + Conversión de modelos

### 4. Requisitos No Funcionales

#### 4.1. Rendimiento

* Latencia: < 200 ms por muestra (inferencia CPU)
* Throughput: > 100 muestras/segundo (batch processing)
* Escalabilidad: Soporte para datasets de hasta 50GB

#### 4.2. Confiabilidad

* Disponibilidad: 99.5% uptime en operaciones críticas
* Manejo de Errores:
  + Validación exhaustiva de inputs
  + Recovery automático de fallos
* Checkpoints: Guardado de estado en entrenamientos largos

#### 4.3. Seguridad

* Autenticación:
  + Variables de entorno para credenciales
  + Exclusión de secrets del repositorio
* Validación:
  + Sanitización de inputs
  + Prevención de inyección SQL
* Control de Acceso:
  + Roles y permisos en bases de datos
  + Encryption de datos sensibles

#### 4.4. Mantenibilidad

* Código:
  + Arquitectura modular
  + Documentación de interfaces
* Versioning:
  + requirements.txt actualizado
  + Control de versiones de modelos
* Testing:
  + Smoke tests básicos
  + Validación cross-platform

#### 4.5. Portabilidad

* Entornos:
  + Soporte multiplataforma (Windows/Linux)
  + Entornos virtuales reproducibles
* Compatibilidad:
  + Backward compatibility de modelos
  + Metadata de características

#### 4.6. Observabilidad

* Logging:
  + Niveles INFO/ERROR/WARNING
  + Formato estructurado
* Métricas:
  + Accuracy, precision, recall
  + Tiempos de ejecución
* Monitoring:
  + Opcional: integración con sistemas de monitoreo

#### 4.7. Cumplimiento Normativo

* Protección de Datos:
  + Análisis de datos sensibles
  + Cumplimiento Ley N° 29733
* Auditoría:
  + Trazabilidad de operaciones
  + Registros de cumplimiento

### 5. Dependencias Críticas

#### 5.1. Software

* Python 3.8+ runtime
* ONNX Runtime 1.22.2+
* Bibliotecas ML (scikit-learn, pandas, numpy)
* Controladores CUDA (si usa GPU)

#### 5.2. Infraestructura

* Acceso a datos de entrenamiento
* Storage para modelos y datasets
* Conectividad de red para updates

#### 5.3. Operacionales

* Personal con conocimientos en:
  + Machine Learning
  + Administración de sistemas
  + Ciberseguridad básica

### 6. Entregables Técnicos

#### 6.1. Código Fuente

* Scripts de entrenamiento y predicción
* Modelos pre-entrenados
* Documentación técnica

#### 6.2. Documentación

* Manuales de instalación
* Guías de uso
* Protocolos de mantenimiento

#### 6.3. Configuración

* + Requirements.txt completo
  + Scripts de deployment
  + Ejemplos de configuración

# Restricciones

### 1. Restricciones Técnicas

* Recursos Limitados: La infraestructura tecnológica existente en algunas instituciones educativas puede ser insuficiente para soportar el sistema, lo que limitará su implementación completa.
* Interoperabilidad: El sistema debe ser compatible con tecnologías y plataformas ya existentes en las instituciones educativas para prevenir problemas de integración.
* Capacidades de Infraestructura: Algunas instituciones pueden no contar con la capacidad adecuada de red y servidores, lo que podría afectar el rendimiento del sistema y su capacidad para procesar datos en tiempo real.

### 2. Restricciones Financieras

* Presupuesto Limitado: Existen limitaciones presupuestarias que pueden afectar la adquisición de hardware, software y otros recursos necesarios para el desarrollo y mantenimiento del sistema.
* Costos de Mantenimiento: La necesidad de prever costos continuos para la actualización y soporte técnico del sistema podría ser una restricción, afectando la sostenibilidad a largo plazo.

### 3. Restricciones Operativas

* Capacitación del Personal: La necesidad de entrenar al personal en el uso del nuevo sistema puede ser una limitación si no se cuenta con el tiempo o los recursos necesarios para la capacitación adecuada.
* Costo y Tiempo de Implementación: Se deben considerar los plazos académicos y otros compromisos de las instituciones que pueden limitar el tiempo disponible para una implementación completa del sistema.

### 4. Restricciones Legales y Regulatorias

* Cumplimiento Normativo: Es imperativo cumplir con las regulaciones locales e internacionales sobre protección de datos personales y privacidad de la información (Ley N° 29733). Esto puede impactar en cómo se manejan y almacenan los datos recogidos por el sistema.
* Consentimiento de Usuarios: La necesidad de obtener el consentimiento de los usuarios para tratar sus datos puede ser una restricción en el desarrollo e implementación del sistema.

### 5. Restricciones Culturales

* Adaptación de la Cultura Organizacional: La resistencia al cambio por parte de los usuarios puede afectar la adopción del sistema en las instituciones educativas, lo que limitara la eficacia del proyecto si no se gestiona adecuadamente.

### 6. Restricciones de Tiempo

* + Cronograma de Implementación: La implementación del sistema podría verse afectada por eventos académicos y otras actividades en las instituciones que restrinjan la disponibilidad del personal y los recursos necesarios.

# Supuestos

El éxito del proyecto se basa en los siguientes supuestos críticos que se han identificado como fundamentales para su ejecución. Estos supuestos representan condiciones que se asumen como verdaderas y que, de no cumplirse, podrían impactar significativamente en el alcance, cronograma o presupuesto del proyecto.

### 8.1. Supuestos Técnicos

Se asume que la infraestructura tecnológica actual de las instituciones educativas de Tacna cumplirá con los requisitos mínimos necesarios para la implementación del sistema. Esto incluye la disponibilidad de conectividad a internet estable con un ancho de banda mínimo de 100 Mbps, equipos de computo con capacidad para ejecutar los componentes del sistema, y entornos virtualizados básicos para las pruebas controladas. Adicionalmente, se supone que las instituciones contarán con personal técnico con conocimientos básicos en administración de sistemas y redes, lo que permitirá una implementación más fluida y una operación adecuada post-implementación.

### 8.2. Supuestos Operativos

El proyecto opera bajo el supuesto de que las instituciones educativas participantes mantendrán un compromiso activo durante todas las fases del proyecto. Esto incluye la disponibilidad del personal clave para las sesiones de capacitación, la provisión de acceso a los sistemas necesarios para las pruebas, y la cooperación en la recolección de datos requeridos para el entrenamiento de los modelos de machine learning. Se asume también que no habrá cambios significativos en la estructura organizacional o prioridades institucionales que afecten la participación comprometida.

### 8.3. Supuestos de Datos

Un supuesto fundamental es la disponibilidad de datos suficientes y de calidad adecuada para el entrenamiento efectivo de los modelos de machine learning. Se asume que se podrá acceder a datos anonimizados de comportamiento de red y procesos que permitan identificar patrones característicos de keyloggers. Además, se supone que estos datos serán representativos de las amenazas reales presentes en el entorno educativo tacneño y que cumplirán con todos los requisitos de protección de datos personales establecidos por la ley peruana.

### 8.4. Supuestos Institucionales

El proyecto assume que las instituciones educativas reconocen la importancia de la ciberseguridad y estarán dispuestas a adoptar los cambios organizacionales necesarios para implementar y mantener el sistema. Esto incluye la asignación de recursos humanos para la operación del sistema, la adaptación de procedimientos existentes, y la integración del sistema con sus procesos actuales de gestión de seguridad. Se asume también que existirá apoyo continuo de la alta dirección institucional.

### 8.5. Supuestos Regulatorios

Se asume que el marco legal peruano actual en materia de protección de datos y ciberseguridad se mantendrá estable durante la ejecución del proyecto, sin cambios regulatorios significativos que requieran modificaciones sustanciales en el diseño o implementación del sistema. Además, se supone que las instituciones contarán con las autorizaciones necesarias para el monitoreo de seguridad dentro de sus redes, en cumplimiento con la Ley N° 29733 y su reglamento.

### 8.6. Supuestos de Entorno

El proyecto assume condiciones ambientales y operativas estables en la región de Tacna durante el período de ejecución. Esto incluye la disponibilidad continua de servicios básicos como energía eléctrica e internet, así como la ausencia de eventos disruptivos mayores que puedan afectar la ejecución del proyecto. Se asume también que las condiciones económicas se mantendrán relativamente estables, sin fluctuaciones significativas que impacten el presupuesto asignado.

### 8.7. Supuestos de Externalidades

Se assume que no surgirán tecnologías disruptivas o amenazas de seguridad completamente nuevas durante el período de desarrollo que hagan obsoleta la aproximación tecnológica seleccionada. Además, se supone que los vendors de software open source mantendrán el soporte de las componentes críticas utilizadas en el proyecto, y que la comunidad de desarrollo continuará proporcionando actualizaciones de seguridad para las dependencias clave.

### 8.8. Supuestos de Sustentabilidad

Finalmente, se asume que las instituciones educativas contarán con los recursos necesarios para dar sostenibilidad al proyecto después del período de implementación inicial. Esto incluye la capacidad de asignar presupuesto para mantenimiento, actualizaciones y capacitación continua del personal. Se supone también que el valor demostrado por el sistema durante la fase piloto justificará la inversión continua en su operación y mejora.

# Resultados Esperados

El proyecto "Sistema web con Machine Learning para la detección anticipada de amenazas informáticas en instituciones educativas de Tacna - 2025" generará resultados significativos en múltiples dimensiones, los cuales se detallan a continuación:

### 1. Resultados de Desempeño Técnico

* Efectividad en la Detección:  
  Se espera alcanzar una tasa de detección del 95% para keyloggers y amenazas similares, superando significativamente los sistemas tradicionales basados en firmas que typically alcanzan entre 60-70% de efectividad.
* Reducción de Falsas Alarmas:  
  El sistema mantendrá una tasa de falsos positivos inferior al 3%, optimizando la eficiencia operativa al minimizar intervenciones innecesarias del personal técnico.
* Tiempos de Respuesta Mejorados:  
  La detección y alerta de amenazas se realizará en menos de 5 segundos desde el inicio de la actividad maliciosa, permitiendo una respuesta oportuna antes del compromiso de datos sensibles.

### 2. Impacto Operacional y Económico

* Eficiencia Operativa:  
  Reducción del 70% en el tiempo dedicado al monitoreo manual de seguridad, permitiendo al personal de TI enfocarse en actividades de mayor valor estratégico.
* Ahorro de Costos:  
  Prevención de pérdidas estimadas en S/ 25,000 anuales por incidentes de seguridad evitados, representando un retorno de inversión (ROI) del 28% anual.
* Optimización de Recursos:  
  Disminución del 60% en los requerimientos de hardware para seguridad mediante la consolidación de funcionalidades en una plataforma unificada.

### 3. Resultados de Capacitación y Desarrollo

* Fortalecimiento de Capacidades:  
  Capacitación de 15+ administradores de sistemas en técnicas avanzadas de machine learning aplicado a ciberseguridad, creando competencias locales especializadas.
* Cultura de Seguridad:  
  Implementación de programas de concienciación que alcanzarán a 200+ docentes y personal administrativo, elevando el nivel general de higiene digital en las instituciones.

### 4. Resultados Institucionales y Estratégicos

* Mejora en la Postura de Seguridad:  
  Implementación de protocolos estandarizados de respuesta a incidentes que reducirán en un 85% el tiempo de mitigación de amenazas.
* Cumplimiento Normativo:  
  Garantía de cumplimiento con la Ley N° 29733 de protección de datos personales, evitando potenciales sanciones de hasta 100 UIT (S/ 495,000).
* Ventaja Competitiva:  
  Posicionamiento de las instituciones participantes como líderes en innovación tecnológica educativa, potentially aumentando la atracción de nuevos estudiantes en un 15%.

### 5. Entregables Tangibles

* Plataforma Tecnológica:  
  Sistema web completo con capacidad de procesamiento de 1000+ eventos por segundo y almacenamiento de 90 días de datos históricos.
* Modelos de Machine Learning:  
  Conjunto de algoritmos entrenados y validados con precision >95%, documentación técnica completa y capacidades de auto-mejora continua.
* Documentación Especializada:  
  Manuales técnicos, guías de usuario y protocolos operativos estandarizados para asegurar la sostenibilidad del proyecto.

### 6. Impacto en la Comunidad Educativa

* Confianza Digital:  
  80% de los usuarios reportarán una mejora significativa en su percepción de seguridad y confianza en los sistemas institucionales.
* Protección de Datos:  
  Salvaguarda de información sensible de 2000+ estudiantes y personal, previniendo potenciales robos de identidad y fraudes.
* Ambiente Educativo Seguro:  
  Creación de un entorno digital protegido que facilite la adopción de tecnologías educativas innovadoras sin comprometer la seguridad.

### 7. Métricas de Evaluación Continua

* Monitorización de Desempeño:  
  Sistema de métricas en tiempo real que trackeará precisión de detección, tiempos de respuesta y disponibilidad del sistema.
* Evaluación de Impacto:  
  Encuestas trimestrales de satisfacción y estudios de caso documentados que demostrarán el valor generado por el sistema.
* Benchmarking:  
  Comparativas periódicas con estándares de la industria y mejores prácticas internacionales en ciberseguridad educativa.

### 8. Sostenibilidad y Escalabilidad

* + Modelo Autosostenible:  
    Framework operativo que permitirá la continuación del proyecto beyond el período de implementación inicial, con costos operativos reducidos en 40%.
  + Capacidad de Expansión:  
    Arquitectura escalable que facilitará la incorporación de 10+ instituciones adicionales sin requerir redesign significativo.
  + Actualizaciones Continuas:  
    Mecanismos de auto-aprendizaje que mantendrán la efectividad del sistema frente a amenazas emergentes y evolucionadas.

# Metodología de implementación

La implementación del "Sistema web con Machine Learning para la detección anticipada de amenazas informáticas en instituciones educativas de Tacna - 2025" seguirá una metodología estructurada en fases iterativas, combinando principios ágiles con prácticas de gestión de proyectos tradicionales para asegurar el éxito del despliegue.

### 1. Fase de Planificación y Análisis

Actividades Clave:

* Realización de un levantamiento exhaustivo de requisitos mediante entrevistas con stakeholders clave (personal administrativo, docentes, equipo de TI)
* Análisis del entorno tecnológico actual de las instituciones educativas participantes
* Definición de métricas de éxito y indicadores clave de desempeño (KPIs)
* Establecimiento del plan de proyecto detallado con cronograma, hitos y asignación de recursos

Entregables:

* Documento de requisitos funcionales y no funcionales
* Matriz de stakeholders y plan de comunicación
* Cronograma detallado con dependencias críticas
* Plan de gestión de riesgos inicial

### 2. Fase de Diseño y Arquitectura

Actividades Clave:

* Diseño de la arquitectura técnica del sistema, incluyendo diagramas de flujo de datos y procesos
* Creación de prototipos de interfaz de usuario para validación con usuarios finales
* Selección de algoritmos de machine learning específicos para detección de keyloggers
* Definición de los estándares de seguridad y protocols de protección de datos

Entregables:

* Documento de arquitectura técnica
* Prototipos de interfaz de usuario aprobados
* Especificaciones técnicas de los modelos de machine learning
* Plan de seguridad y privacidad de datos

### 3. Fase de Desarrollo Iterativo

Actividades Clave:

* Desarrollo en sprints de dos semanas siguiendo metodología Scrum
* Implementación modular del sistema con integración continua
* Desarrollo paralelo de los componentes de frontend, backend y modelos de ML
* Revisiones técnicas semanales para asegurar la calidad del código

Entregables:

* Componentes desarrollados y probados unitariamente
* Repositorio de código con control de versiones
* Documentación técnica de desarrollo
* Pipeline de integración continua configurado

### 4. Fase de Pruebas y Validación

Actividades Clave:

* Ejecución de pruebas unitarias, de integración y de sistema
* Simulación de ataques controlados para validar la efectividad de la detección
* Pruebas de carga y estrés para verificar el rendimiento del sistema
* Auditoría de seguridad independiente del código y la configuración

Entregables:

* Reporte completo de pruebas con resultados
* Matriz de trazabilidad de requisitos-pruebas
* Documentación de vulnerabilidades identificadas y remediadas
* Certificación de cumplimiento de seguridad

### 5. Fase de Implementación Piloto

Actividades Clave:

* Despliegue gradual en las tres instituciones educativas piloto
* Migración y configuración de datos iniciales
* Monitoreo intensivo durante las primeras 72 horas post-implementación
* Establecimiento de canales de soporte técnico prioritario

Entregables:

* Sistema implementado en ambiente productivo
* Documentación de configuración específica por institución
* Plan de rollback en caso de contingencias
* Reporte inicial de operación post-implementación

### 6. Fase de Capacitación y Transferencia de Conocimiento

Actividades Clave:

* Desarrollo de programas de capacitación diferenciados por perfiles de usuario
* Sesiones de entrenamiento práctico para administradores del sistema
* Talleres de concienciación en ciberseguridad para la comunidad educativa
* Transferencia de conocimiento al personal técnico local

Entregables:

* Materiales de capacitación (manuales, videos, guías rápidas)
* Registros de asistencia y evaluación de participantes
* Plan de sostenibilidad y conocimiento institucional
* Certificados de competencia para usuarios avanzados

### 7. Fase de Monitoreo y Optimización Continua

Actividades Clave:

* Monitoreo continuo del desempeño del sistema mediante dashboards
* Recolección sistemática de feedback de usuarios
* Análisis periódico de métricas de efectividad y rendimiento
* Implementación de mejoras basadas en datos reales de uso

Entregables:

* Reportes periódicos de desempeño del sistema
* Registro de incidentes y acciones de mejora
* Actualizaciones incrementales del sistema
* Plan de mejora continua a largo plazo

### 8. Fase de Escalamiento y Expansion

Actividades Clave:

* Evaluación de resultados obtenidos en la fase piloto
* Adaptación del sistema para instituciones adicionales
* Desarrollo de documentación para escalamiento
* Establecimiento de modelos de soporte para nuevas instituciones

Entregables:

* Modelo de escalamiento documentado
* Kit de implementación para nuevas instituciones
* Plan de expansión regional
* Acuerdos de soporte y mantenimiento a largo plazo

### Enfoque de Gestión

Gestión de Riesgos:

* Identificación temprana de riesgos técnicos y operativos
* Planes de mitigación para riesgos de alta probabilidad e impacto
* Revisión mensual de la matriz de riesgos

Comunicación:

* Reportes semanales de progreso a stakeholders
* Comités de seguimiento quincenales con instituciones participantes
* Canal de comunicación abierto para issues críticos

Control de Calidad:

* + Revisión por pares de todo el código desarrollado
  + Auditorías regulares de cumplimiento de estándares
  + Mediciones objetivas de satisfacción de usuarios

# Actores claves

| **Actor** | **Rol** |
| --- | --- |
| **Sebastián Arce Bracamonte** | Investigador Principal; responsable de la supervisión general y ejecución del proyecto. |
| **Renzo Taco Coayla** | Asesor Académico; proporciona orientación y revisión al proyecto. |
| **Equipo de Desarrollo de Software** | Especialistas en programación; responsables del desarrollo y ajustes del sistema. |
| **Personal de Seguridad Informática** | Encargados de la implementación y monitoreo de la seguridad del sistema en las instituciones educativas. |
| **Administradores de la Institución** | Proporcionan soporte en la adopción del sistema y garantizan la alineación con las políticas institucionales. |
| **Docentes y Estudiantes** | Usuarios finales; serán beneficiarios del sistema, proporcionando retroalimentación sobre su efectividad. |
| **Personal de TI** | Encargados del mantenimiento y soporte técnico del sistema, así como de la capacitación del personal. |
| **Evaluadores Externos** | Especialistas en ciberseguridad que realizarán auditorías y evaluaciones del sistema post-implementación. |

# Papel y responsabilidad personal

### Investigador Principal (Sebastián Arce Bracamonte)

* Liderazgo Técnico: Diseño de arquitectura de machine learning y supervisión del desarrollo
* Gestión de Proyecto: Control del cronograma y presupuesto asignado
* Coordinación: Función como enlace principal entre la universidad y las instituciones piloto
* Entregables: Modelos de machine learning con alta precisión y documentación técnica completa
* Métrica: Cumplimiento de hitos y alta satisfacción de stakeholders

### Asesor Académico (Renzo Taco Coayla)

* Validación Técnica: Revisión de algoritmos y enfoques de machine learning
* Alineamiento Educativo: Garantizar la relevancia pedagógica del sistema
* Transferencia de Conocimiento: Capacitación en técnicas avanzadas
* Entregables: Protocolos de validación y programas de capacitación
* Métrica: Cumplimiento de estándares académicos

### Equipo de Desarrollo

* Implementación: Desarrollo de componentes frontend, backend y machine learning
* Calidad: Producción de código con amplia cobertura de pruebas
* Documentación: Elaboración de manuales técnicos y de usuario
* Entregables: Sistema web funcional con APIs documentadas
* Métrica: Mínimos bugs críticos post-implementación

### Personal de Seguridad Informática

* Operación: Monitoreo continuo durante la fase piloto
* Respuesta: Mitigación rápida de incidentes de seguridad
* Capacitación: Entrenamiento a administradores institucionales
* Entregables: Protocolos de respuesta y reportes de incidentes
* Métrica: Reducción significativa de incidentes de seguridad

### Administradores Institucionales

* Recursos: Provisión de infraestructura tecnológica requerida
* Adopción: Facilitación de la implementación en entornos reales
* Sostenibilidad: Planificación del mantenimiento post-proyecto
* Entregables: Acuerdos institucionales y recursos asignados
* Métrica: Implementación exitosa en todas las instituciones

### Docentes y Estudiantes

* Feedback: Retroalimentación continua sobre usabilidad del sistema
* Adopción: Participación activa en capacitaciones y pruebas
* Cultura de Seguridad: Aplicación de mejores prácticas aprendidas
* Entregables: Encuestas de satisfacción y casos de uso
* Métrica: Alta satisfacción con la usabilidad del sistema

### Personal de TI Institucional

* Soporte: Mantenimiento diario del sistema implementado
* Actualizaciones: Aplicación de parches y mejoras necesarias
* Monitoreo: Reporte proactivo de issues y problemas
* Entregables: Tickets de soporte y reportes de desempeño
* Métrica: Alta disponibilidad del sistema en producción

### Evaluadores Externos

* + Auditoría: Evaluación independiente de la efectividad del sistema
  + Benchmarking: Comparación con estándares de la industria
  + Recomendaciones: Elaboración de planes de mejora continua
  + Entregables: Reportes de auditoría y recomendaciones
  + Métrica: Objetividad total en las evaluaciones realizadas

# Plan de monitoreo y evaluación

| **Componente** | **Actividades** | **Frecuencia** | **Responsables** | **Entregables** |
| --- | --- | --- | --- | --- |
| **Monitoreo Inicial** | Establecimiento de línea base de seguridad | Inicio del proyecto | Equipo de Seguridad | Reporte de estado inicial de ciberseguridad |
|  | Capacitación inicial del personal | Mes 1 | Asesor del Proyecto | Registros de capacitación y evaluaciones |
| **Monitoreo Continuo** | Revisión de alertas y eventos de seguridad | Diario | Equipo de Seguridad | Logs de eventos y alertas |
|  | Actualización de modelos de Machine Learning | Mensual | Equipo de Desarrollo | Modelos actualizados y reportes de rendimiento |
| **Evaluación de Desempeño** | Medición de tasa de detección y tiempo de respuesta | Semanal | Equipo de Seguridad | Métricas de desempeño del sistema |
|  | Encuestas de satisfacción de usuarios | Trimestral | Asesor del Proyecto | Reportes de satisfacción usuaria |
| **Evaluación de Impacto** | Análisis comparativo de incidentes (pre-post implementación) | Mensual | Equipo de Seguridad | Reportes de reducción de incidentes |
|  | Cálculo de ROI y beneficios del sistema | Anual | Asesor del Proyecto | Análisis de retorno de inversión |
| **Mejora Continua** | Revisión de falsos positivos/negativos | Mensual | Equipo de Desarrollo | Ajustes a modelos y algoritmos |
|  | Actualización de funcionalidades del sistema | Trimestral | Equipo de Desarrollo | Nuevas versiones del sistema |
| **Comunicación** | Reportes de estado del sistema | Semanal | Equipo de Seguridad | Dashboards y reportes ejecutivos |
|  | Reuniones de revisión con stakeholders | Mensual | Asesor del Proyecto | Minutas y plan de acción |
| **Aseguramiento de Calidad** | Auditorías de seguridad del sistema | Semestral | Evaluadores Externos | Reportes de auditoría y recomendaciones |
|  | Pruebas de penetración controladas | Anual | Equipo de Seguridad | Reportes de vulnerabilidades y remediación |
| **Consideraciones Éticas** | Verificación de cumplimiento de Ley N° 29733 | Continuo | Asesor del Proyecto | Certificaciones de cumplimiento |
|  | Revisión de políticas de privacidad y consentimiento | Trimestral | Equipo de Seguridad | Documentación actualizada de políticas |

## Indicadores Clave de Desempeño

| Indicador | Meta | Frecuencia de Medición | Responsable |
| --- | --- | --- | --- |
| Tasa de detección de amenazas | >95% | Semanal | Equipo de Seguridad |
| Tiempo de respuesta | <5 minutos | Diario | Equipo de Seguridad |
| Satisfacción del usuario | >90% | Trimestral | Asesor del Proyecto |
| Reducción de incidentes | 85% | Mensual | Equipo de Seguridad |
| Disponibilidad del sistema | 99.5% | Continuo | Equipo de Desarrollo |
| Falsos positivos | <3% | Mensual | Equipo de Desarrollo |

## Herramientas de Monitoreo

| Tipo | Herramientas |
| --- | --- |
| Monitoreo Técnico | ELK Stack, Prometheus/Grafana, Wireshark |
| Gestión de Incidentes | Jira Service Desk, Sistema de ticketing institucional |
| Evaluación de Usuario | Encuestas Online, Formularios Google, Entrevistas estructuradas |
| Análisis de Datos | Python Pandas, Jupyter Notebooks, Power BI |
| Comunicación | Slack, Microsoft Teams, Correo institucional |

# Cronograma del proyecto

| **Fase** | **Mes** | **Actividades Clave** | **Entregables** | **Responsable** |
| --- | --- | --- | --- | --- |
| **Fase 1: Preparación** | Agosto | - Análisis de infraestructura existente  - Reclutamiento de personal  - Definición de métricas iniciales  - Adquisición de hardware/software | - Informe de diagnóstico  - Plan de recursos humanos  - Línea base de seguridad | Investigador Principal |
| **Fase 2: Desarrollo** | Septiembre-Octubre | - Desarrollo de algoritmos ML  - Implementación del frontend  - Creación de APIs  - Integración de módulos | - Modelos ML entrenados  - Sistema web funcional  - Documentación técnica | Equipo de Desarrollo |
| **Fase 3: Validación** | Noviembre | - Pruebas de penetración  - Validación de modelos  - Simulación de ataques  - Optimización del sistema | - Reporte de seguridad  - Certificación de calidad  - Manual de operaciones | Equipo de Seguridad |
| **Fase 4: Implementación** | Diciembre | - Despliegue en instituciones piloto  - Capacitación de usuarios  - Monitoreo inicial  - Ajustes finales | - Sistema operativo  - Registros de capacitación  - Reporte de implementación | Todos los equipos |

# Hitos de entregable

| Hito | Fecha | Descripción | Entregable |
| --- | --- | --- | --- |
| H1 | 15/08/2025 | Finalización análisis inicial | Informe diagnóstico completo |
| H2 | 30/09/2025 | Completación desarrollo modelos ML | Modelos con >90% precisión |
| H3 | 31/10/2025 | Sistema integrado funcionando | Demo funcional completa |
| H4 | 15/11/2025 | Validación de seguridad | Certificado de pruebas de penetración |
| H5 | 15/12/2025 | Implementación en primeras 3 instituciones | Reporte de implementación exitosa |

# **RESUMEN EJECUTIVO**

| **Aspecto** | **Descripción** |
| --- | --- |
| **Nombre del Proyecto** | Sistema web con Machine Learning para la detección anticipada de keyloggers en instituciones educativas de Tacna - 2025 |
| **Propósito del Proyecto** | Implementar un sistema inteligente de detección proactiva de keyloggers mediante técnicas de Machine Learning para fortalecer la ciberseguridad en instituciones educativas |
| **Resultados Esperados** | - Reducción del 85% en incidentes de seguridad  - Detección del 95% de amenazas en menos de 5 segundos  - Ahorro estimado de S/ 25,000 anuales en costos de remediación  - Mejora del 90% en percepción de seguridad |
| **Población Objetivo** | 3 instituciones educativas piloto en Tacna, beneficiando a 2,000+ usuarios entre estudiantes, docentes y personal administrativo |
| **Monto de Inversión** | S/ 101,860 |
| **Duración del Proyecto** | 6 meses (Agosto - Diciembre 2025) |
| **Retorno de Inversión** | 28% anual, con recuperación de la inversión en 2.1 años |
| **Sostenibilidad** | Modelo autosostenible mediante ahorros generados, con capacidad de expansión a más instituciones |
| **Alineamiento Estratégico** | Cumplimiento de Ley N° 29733 de protección de datos y Plan Nacional de Ciberseguridad 2023-2026 |