**TALENTO TECH 2024-MINTIC**

**FORMATO DE PRESENTACIÓN “PLAN DE PROYECTO TI”**

**Contexto específico de aplicación del proyecto** *(Marque con una X)*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **AGRO** | **EDUCACIÓN** | **TURISMO** | **GOBIERNO** | **FINANZAS** | **MARKETING** | **SALUD** | **OTRO** |
|  |  |  |  |  |  |  | **X** |

**Cohorte #: \_\_\_\_ Año: 2024 Tutor: Andrés Felipe Escallon**

**Nombre del Proyecto (y del producto/servicio):**

|  |
| --- |
| Predicción de PQRS en Departamentos Usando Series de Tiempo y Redes Neuronales |

**Departamento de residencia del estudiante:**

|  |
| --- |
| Bogotá D.C |

**Municipio de residencia del estudiante:**

|  |
| --- |
| Bogotá D.C |

**Rural:** (Marque con una X)

|  |  |  |  |
| --- | --- | --- | --- |
| **SI** |  | **NO** | **X** |
| **Vereda o Corregimiento:** | | | | | N/A |

**Autor (es):**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Nombres y Apellidos** | **Tipo de identificación** | **No. identificación** | **Curso: Programación, Inteligencia Artificial, Análisis Datos, Block Chain, Arquitectura Nube** | **Nivel: Explorador, Integrador, Innovador** | **Modalidad: Virtual, Semipresencial o Presencial** |
|  | William Raúl Mora Valencia | C.C | x.xxx.xxx.xxx | Inteligencia Artificial | Integrador | Virtual |

**Palabras clave:**

|  |  |
| --- | --- |
| **Palabra clave 1** | Predicción |
| **Palabra clave 2** | Series de Tiempo |
| **Palabra clave 3** | Redes neuronales |
| **Palabra clave 4** | PQRSD |

**Planteamiento del problema que solucionará el producto/servicio:**

|  |
| --- |
| En diversas entidades gubernamentales y privadas, las PQRS (Peticiones, Quejas, Reclamos, Sugerencias y Denuncias) son un indicador clave del nivel de satisfacción de los ciudadanos y usuarios respecto a los servicios ofrecidos. La gestión eficiente de estos mecanismos no solo mejora la percepción de los ciudadanos, sino que también ayuda a optimizar procesos internos e identificar áreas de mejora.  Sin embargo, la alta variabilidad en la cantidad y el tipo de PQRS recibidas en diferentes departamentos hace que su gestión y prevención sea un reto. Factores como cambios estacionales, decisiones administrativas, eventos socioeconómicos y la disponibilidad de servicios pueden afectar significativamente el volumen de PQRS que una entidad recibe en un determinado periodo de tiempo.  Actualmente, muchas entidades reaccionan ante estos picos de solicitudes y reclamos sin herramientas predictivas adecuadas, lo que puede generar demoras en las respuestas y en la implementación de soluciones. Esto podría traducirse en un impacto negativo en la relación entre la ciudadanía y las entidades. |

**Pertinencia del proyecto TI:**

|  |
| --- |
| La transformación digital y la creciente demanda de eficiencia en la gestión pública han impulsado la necesidad de soluciones basadas en datos para mejorar la calidad del servicio ofrecido a la ciudadanía. En este contexto, la implementación de modelos predictivos para anticipar el volumen de PQRS (Peticiones, Quejas, Reclamos, Sugerencias y Denuncias) en los diferentes departamentos del país responde a una necesidad crítica de planificación y mejora continua.  **Mercado:**  El mercado objetivo para la implementación de un modelo predictivo basado en series temporales y redes neuronales para predecir el volumen de PQRS es amplio, y abarca tanto el sector público como el privado. Este proyecto tiene el potencial de beneficiar a cualquier entidad que gestione una cantidad considerable de interacciones o solicitudes ciudadanas, ya que su objetivo es optimizar la respuesta a peticiones y mejorar la planificación de recursos. |

**Estado del Arte de productos/servicios existentes y ventajas comparativas:**

En la revisión realizada por el momento no se encontró ningún servicio existente, actualmente los enfoques usados para los PQRSD son hacia la clasificación.

**Marco Legal y Ético**

|  |
| --- |
| **Protección de Datos Personales:**   * **Ley 1581 de 2012:** Esta ley regula la protección de datos personales en Colombia. Establece principios para la recolección, almacenamiento, uso y transferencia de datos personales. * **Política de Privacidad:** contar con una política de privacidad clara que explique cómo se recopilan, usan y protegen los datos personales.   **Regulaciones de Seguridad:**   * Implementar medidas adecuadas para garantizar la seguridad de los datos, incluyendo cifrado y controles de acceso. La Ley 1581 también requiere medidas para proteger los datos contra accesos no autorizados y pérdidas.   **Confidencialidad:**   * Proteger la confidencialidad de la información y evitar la divulgación no autorizada. Solo las personas autorizadas tendrán acceso a datos sensibles. |

**ANÁLISIS DE RIESGOS:**

### 1. **Riesgos de Datos**

* **Riesgo:** **Acceso no autorizado** a datos personales o sensibles.
  + **Mitigación:** Implementa medidas de seguridad robustas como cifrado, autenticación de usuarios y controles de acceso estrictos.
* **Riesgo:** **Pérdida o corrupción de datos.**
  + **Mitigación:** Realiza copias de seguridad periódicas y asegúrate de tener un plan de recuperación ante desastres.
* **Riesgo:** **Datos incompletos o incorrectos.**
  + **Mitigación:** Realiza una validación exhaustiva de los datos antes del análisis y establece procedimientos para corregir errores.

### 2. **Riesgos Legales**

* **Riesgo:** **Incumplimiento de la Ley de Protección de Datos Personales (Ley 1581 de 2012).**
  + **Mitigación:** Asegúrate de obtener el consentimiento informado de los titulares de los datos y de cumplir con todos los requisitos legales de protección y privacidad.
* **Riesgo:** **Uso indebido de datos personales.**
  + **Mitigación:** Define claramente los límites de uso de los datos en tu política de privacidad y garantiza que solo se usen para los fines previstos.

### 3. **Riesgos Éticos**

* **Riesgo:** **Discriminación o sesgo en las predicciones.**
  + **Mitigación:** Evalúa y ajusta el modelo para detectar y minimizar sesgos. Realiza auditorías periódicas para asegurar la equidad.
* **Riesgo:** **Impacto negativo en los individuos o comunidades.**
  + **Mitigación:** Realiza evaluaciones de impacto y consulta con partes interesadas para entender y mitigar cualquier efecto adverso potencial.

### 4. **Riesgos Técnicos**

* **Riesgo:** **Desempeño deficiente del modelo.**
  + **Mitigación:** Realiza pruebas exhaustivas y ajusta los hiperparámetros del modelo. Monitorea el desempeño y ajusta según sea necesario.
* **Riesgo:** **Falta de escalabilidad.**
  + **Mitigación:** Diseña el sistema para ser escalable y realiza pruebas de carga para asegurar que puede manejar el volumen de datos y consultas.

**Objetivos:**

|  |
| --- |
| **Generales**  **Desarrollar un Modelo Predictivo:**   * **Objetivo:** Crear y entrenar un modelo de redes neuronales capaz de predecir con precisión el volumen de PQRS en diferentes departamentos basándose en datos históricos de series temporales.   **Mejorar la Gestión de PQRS:**   * **Objetivo:** Utilizar el modelo predictivo para optimizar la gestión y respuesta a PQRS en diferentes departamentos, permitiendo una asignación más eficiente de recursos.   **Específicos**  **Preparación de Datos:**   * **Objetivo:** Recopilar, limpiar y preprocesar los datos históricos de PQRS de manera que sean adecuados para el análisis de series temporales.   **Desarrollo del Modelo:**   * **Objetivo:** Implementar y entrenar diferentes arquitecturas de redes neuronales recurrentes (RNN), como LSTM o GRU, para la predicción de PQRS.   **Evaluación del Modelo:**   * **Objetivo:** Evaluar el desempeño del modelo utilizando métricas de rendimiento como MAE, RMSE y MAPE. |

**Metodología**:

|  |
| --- |
| **metodología basada en fases:**   Definir claramente el problema y los objetivos.   Recopilar y preparar datos de PQRS por departamento.   Desarrollar un modelo predictivo basado en redes neuronales (LSTM, GRU).   Evaluar y validar el modelo en datos no vistos.   Implementar el modelo en producción. |

**Plazo:** Duración del proyecto.

|  |  |
| --- | --- |
| **SEMANAS** | **DIAS** |
| 10 |  |

**CRONOGRAMA DE ACTIVIDADES** (Diagrama de Gantt):

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Actividad** | **S1** | **S2** | **S3** | **S4** | **S5** | **S**  **6** | **S**  **7** | **S**  **8** | **S**  **9** | **S**  **10** | **Responsable** |
| 1 | Definición clara del problema y los objetivos del proyecto. | X |  |  |  |  |  |  |  |  |  | William Mora |
| 2 | Recopilación de datos históricos de PQRS. |  | X |  |  |  |  |  |  |  |  |  |
| 3 | Limpieza de datos (manejo de valores faltantes, corrección de errores, etc.). |  |  | X |  |  |  |  |  |  |  |  |
| 4 | Estructuración de los datos en series temporales. |  |  |  | X |  |  |  |  |  |  |  |
| 5 | Creación de diccionarios con los datos de entrenamiento y prueba, manteniendo la integridad temporal. |  |  |  | X |  |  |  |  |  |  |  |
| 6 | Normalización y transformación de los datos para el modelo. |  |  |  |  | X |  |  |  |  |  |  |
| 7 | Selección de arquitecturas de redes neuronales (LSTM, GRU, etc.). |  |  |  |  |  | X |  |  |  |  |  |
| 8 | Definición de la estructura del modelo en Keras y TensorFlow. |  |  |  |  |  |  | X |  |  |  |  |
| 9 | Entrenamiento inicial del modelo con los datos de entrenamiento. |  |  |  |  |  |  | X |  |  |  |  |
| 10 | Selección del mejor modelo basado en las métricas de evaluación. |  |  |  |  |  |  |  | X |  |  |  |
| 11 | Evaluación completa del modelo utilizando métricas como MAE, RMSE, MAPE. |  |  |  |  |  |  |  | X |  |  |  |
| 12 | Validación cruzada y pruebas adicionales para asegurar la robustez del modelo. |  |  |  |  |  |  |  |  | X |  |  |
| 13 | Validación final con datos reales o simulados de PQRS recientes. |  |  |  |  |  |  |  |  |  | X |  |

**PRESUPUESTO:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **PRESUPUESTO** | | | | | | | |
|  | | **MANO DE OBRA** | | **MATERIALES** | | **FIJOS** | **BALANCE** |
| **TAREA** | **DESCRIPCIÓN** | **HORA** | **$/HR** | **UNIDAD** | **$/UNIDAD** | **ALQUILER** | **GASTADO** |
| Equipo | HP ProBook 440 G8 | 0,0 | $0 | $1 | $3.500.000 | $0 | $3.500.000 |
| Definición del Proyecto | Definición clara del problema y los objetivos del proyecto. | 20,0 | $150.000 | $0 | $0 | $0 | $3.000.000 |
| Preparación de los Datos | Recopilación de datos históricos de PQRS. | 20,0 | $150.000 | $0 | $0 | $0 | $3.000.000 |
| Preparación de los Datos | Limpieza de datos (manejo de valores faltantes, corrección de errores, etc.). | 20,0 | $150.000 | $0 | $0 | $0 | $3.000.000 |
| Preparación de los Datos | Estructuración de los datos en series temporales. | 10,0 | $150.000 | $0 | $0 | $0 | $1.500.000 |
| Preparación de los Datos | Creación de diccionarios con los datos de entrenamiento y prueba, manteniendo la integridad temporal. | 10,0 | $150.000 | $0 | $0 | $0 | $1.500.000 |
| Preparación de los Datos | Normalización y transformación de los datos para el modelo. | 20,0 | $150.000 | $0 | $0 | $0 | $3.000.000 |
| Almacenamiento de Datos | Servicios AWS | 0,0 | $0 | $0 | $0 | $700.000 | $700.000 |
| Desarrollo del Modelo | Selección de arquitecturas de redes neuronales (LSTM, GRU, etc.). | 20,0 | $150.000 | $0 | $0 | $0 | $3.000.000 |
| Desarrollo del Modelo | Definición de la estructura del modelo en Keras y TensorFlow. | 10,0 | $150.000 | $0 | $0 | $0 | $1.500.000 |
| Desarrollo del Modelo | Entrenamiento inicial del modelo con los datos de entrenamiento. | 10,0 | $150.000 | $0 | $0 | $0 | $1.500.000 |
| Desarrollo del Modelo | Selección del mejor modelo basado en las métricas de evaluación. | 5,0 | $150.000 | $0 | $0 | $0 | $750.000 |
| Evaluación y Validación | Evaluación completa del modelo utilizando métricas como MAE, RMSE, MAPE. | 15,0 | $150.000 | $0 | $0 | $0 | $2.250.000 |
| Evaluación y Validación | Validación cruzada y pruebas adicionales para asegurar la robustez del modelo. | 20,0 | $150.000 | $0 | $0 | $0 | $3.000.000 |
| Almacenamiento de Datos | Servicios AWS | 0,0 | $0 | $0 | $0 | $700.000 | $700.000 |
| Evaluación y Validación | Validación final con datos reales o simulados de PQRS recientes. | 20,0 | $150.000 | $0 | $0 | $0 | $3.000.000 |
| **SUBTOTAL** |  | **200** | **$30.000.000** | **0,0** | **$3.500.000** | **$1.400.000** | **$34.900.000** |