



UNIVERSIDAD TECNOLÓGICA INDOAMÉRICA

**FACULTAD DE INGENIERÍAS EN TECNOLOGÍAS DE LA
INFORMACIÓN Y COMUNICACIÓN
CARRERA DE INGENIERÍA EN SISTEMAS**

TEMA:

**“DESARROLLO DE UN SISTEMA DE PREDICCIÓN DE CONTAGIOS Y
CERCO EPIDEMIOLÓGICO DEL COVID-19 EN LA PROVINCIA DE
TUNGURAHUA PARA EL HONORABLE GOBIERNO PROVINCIAL DE
TUNGURAHUA.”**

Trabajo de titulación previo a la obtención del título de ingeniero en sistemas

Autor

Llerena Ortiz Kevin Estalin

Tutor

Ing. Franklin Adrián Castillo Ledesma MSc.

Ambato – Ecuador

2022

AUTORIZACIÓN PARA REPOSITORIO DIGITAL

Yo, Llerena Ortiz Kevin Estalin, declaro ser autor del Trabajo de Titulación con el tema “DESARROLLO DE UN SISTEMA DE PREDICCIÓN DE CONTAGIOS Y CERCO EPIDEMIOLÓGICO DEL COVID-19 EN LA PROVINCIA DE TUNGURAHUA PARA EL HONORABLE GOBIERNO PROVINCIAL DE TUNGURAHUA.”, como requisito para optar al grado de “Ingeniero en Sistemas” y autorizo al Sistema de Bibliotecas de la Universidad Tecnológica Indoamérica, para que con fines netamente académicos divulgue esta obra a través del Repositorio Digital Institucional (RDI-UTI).

Los usuarios del RDI-UTI podrán consultar el contenido de este trabajo en las redes de información del país y del exterior, con las cuales la Universidad tenga convenios. La Universidad Tecnológica Indoamérica no se hace responsable por el plagio o copia del contenido parcial o total de este trabajo.

Del mismo modo, acepto que los Derechos de Autor, Morales y Patrimoniales, sobre esta obra, serán compartidos entre mi persona y la Universidad Tecnológica Indoamérica, y que no tramitaré la publicación de esta obra en ningún otro medio, sin autorización expresa de la misma. En caso de que exista el potencial de generación de beneficios económicos o patentes, producto de este trabajo, acepto que se deberán firmar convenios específicos adicionales, donde se acuerden los términos de adjudicación de dichos beneficios.

Para constancia de esta autorización, en la ciudad de Ambato, a los 14 días del mes de febrero de 2022, firmo conforme:

Autor: Llerena Ortiz Kevin Estalin

Firma: 

Número de Cédula: 160069622-1

Dirección: 10 de Agosto y Lucindo Ortega

Correo Electrónico: kevinllerena1995@gmail.com

Teléfono: +593 992808735

APROBACION DEL ASESOR

En mi calidad de catedrático Asesor del Proyecto de grado previo a la obtención del título de Ingeniero en Sistemas, titulado “Desarrollo de un sistema de predicción de contagios y cerco epidemiológico del COVID-19 en la provincia de Tungurahua para el Honorable Gobierno Provincial de Tungurahua.”, elaborado por el señor estudiante: Kevin Llerena. Certifico que dicho proyecto ha sido revisado en todas sus partes y considero que reúne los requisitos y méritos suficientes para ser sometido a la presentación pública y evaluación por parte del tribunal examinador que se designe.

Ambato, 14 de febrero de 2022



Ing. Franklin Castillo
ASESOR

DECLARACION DE AUTENTICIDAD

El abajo firmante, declara que los contenidos y resultados obtenidos en el presente proyecto, como requerimiento previo para la obtención del título de Ingeniero en Sistemas, son absolutamente originales, auténticos, personales y de exclusiva responsabilidad legal y académica del autor.



Kevin Estalin Llerena Ortiz

C.I: 1600696221

APROBACION DEL TRIBUNAL

Los Miembros del Tribunal Examinador designado por la Comisión Académica aprueban el trabajo de titulación cuyo tema es: “Desarrollo de un sistema de predicción de contagios y cerco epidemiológico del COVID-19 en la provincia de Tungurahua para el Honorable Gobierno Provincial de Tungurahua.”, de acuerdo con las disposiciones reglamentarias emitidas por la Universidad Tecnológica Indoamérica para la obtención del Título de Ingeniero en Sistemas. Por lo tanto, autorizamos al postulante para su impresión, empastado y sustentación pública.

Ambato, 14 de febrero de 2022



Firmado electrónicamente por:
**MARIO ERNESTO
MIRANDA SANCHEZ**

Ing. Mario Ernesto Miranda Sánchez

PRESIDENTE



Firmado electrónicamente por:
**WILSON PATRICIO
PEÑAHERRERA
ACURIO**

Ing. Wilson Patricio Peñaherrera Acurio

VOCAL 1



Firmado electrónicamente por:
**PATRICIO
GUSTAVO LARA
ALVAREZ**

Ing. Patricio Gustavo Lara Alvares

VOCAL 2

AGRADECIMIENTO

Innumerables personas apoyaron mi esfuerzo en esta tesis. Los profesores Franklin Castillo, Wilson Peñaherrera, Mario Miranda proporcionaron comentarios invaluable sobre mi análisis y encuadre, a veces respondiendo correos electrónicos a altas horas de la noche y temprano en la mañana. Quiero agradecer especialmente al profesor Cesar Guevara y Galo Robayo, quienes me encaminaron hacia el tema del Machine learning, y su aliento para escribir esta tesis me hizo confiar en que el tema era digno de mi investigación y que mi investigación era digna del tema.

También estoy en deuda con varios estudiantes de la Universidad Tecnológica Indoamérica, me ayudaron a encontrar mi equilibrio cuando comencé este proceso.

Por último, mi familia merece una gratitud infinita: mi padre por enseñarme a apreciar que el esfuerzo tiene su recompensa, mi madre por enseñarme a escribir con un propósito conciso y mi pareja por enseñarme que una afirmación de un tema no es necesariamente algo malo. A mi familia le doy todo, incluido esto.

Kevin Estalin Llerena Ortiz

DEDICATORIA

El presente trabajo de tesis lo dedico principalmente a Dios, por ser el inspirador y darnos fuerza para continuar en este proceso de obtener uno de los anhelos más deseados, a mis padres Fausto Llerena y Yolanda Ortiz, por su amor, trabajo y sacrificio en todos estos años especialmente el factor económico brindándome todo lo necesario para poder estudiar fuera de mi ciudad residencial que es el mejor regalo que recibí y por eso les dedico este trabajo, quienes también me enseñaron que el esfuerzo tiene sus recompensas y aunque tarde siempre llegan; los amo!.

A todos mis compañeros que se con certeza que todos nos dedicamos mutuamente este logro, por todo lo que vivimos en el proceso.

Al Honorable Gobierno Provincial de Tungurahua que me acogió para realizar este proyecto que estoy seguro será de gran utilidad para el pueblo.

Kevin Estalin Llerena Ortiz

RESUMEN

UNIVERSIDAD TECNOLÓGICA INDOAMÉRICA

FACULTAD DE INGENIERÍA EN SISTEMAS

TEMA

“Desarrollo de un sistema de predicción de contagios y cerco epidemiológico del COVID-19 en la provincia de Tungurahua para el honorable gobierno provincial de Tungurahua”

AUTOR: KEVIN ESTALIN LLERENA ORTIZ

ASESOR: ING FRANKLIN CASTILLO

RESUMEN

La pandemia de covid-19 a nivel mundial tuvo un impacto muy negativo en la población del país Ecuador particularmente nos referimos a la provincia de Tungurahua en muchas áreas como salud, económica y social dando como resultado la catástrofe social que conlleva atravesar una pandemia. En la provincia de Tungurahua será implementado un sistema integral que ayudara a las autoridades pertinentes en la toma de decisiones, se implantará un algoritmo de predicción desarrollado en Python para predecir la curva de contagios de covid-19 utilizando Machine learning en la provincia de Tungurahua y así alertar a las autoridades, también se desarrollara un mapa de calor elaborado en QGIS para su posterior uso en cercos epidemiológicos que ayudaran a la ciudadanía mantenerse al margen de ciertos sectores de la provincia; todo esto será integrado en una aplicación web programada con las tecnologías Python, Javascript, Html, Css para su fácil manejo y lectura, siendo intuitivo para el usuario final; los datos recolectados para el funcionamiento del sistema serán otorgados por el Ministerio de salud. La implementación de este sistema tiene como objetivo principal evitar el colapso del sistema de salud y así reducir la tasa de mortalidad en la población.

Descriptores: Datos – Machine learning – Mapas de calor.

ABSTRACT
TECHNOLOGICAL UNIVERSITY INDOAMERICA

SUBJECT

“Development of a system for predicting infections and epidemiological encirclement of COVID-19 in the province of Tungurahua for the honorable provincial government of Tungurahua”

AUTHOR: KEVIN ESTALIN LLERENA ORTIZ

TUTOR: ING FRANKLIN CASTILLO

SUMMARY

The covid-19 pandemic worldwide had a very negative impact on the population of the country Ecuador particularly we refer to the province of Tungurahua in many areas such as health, economic and social resulting in the social catastrophe that entails going through a pandemic. In the province of Tungurahua, a comprehensive system will be implemented to help the relevant authorities in decision making, a prediction algorithm developed in Python will be implemented to predict the curve of covid-19 infections using Machine Learning in the province of Tungurahua and thus alert the authorities, a heat map prepared in QGIS will also be developed for later use in epidemiological fences that will help citizens maintain themselves outside certain sectors of the province; all this will be integrated into a web application programmed with Python, JavaScript, Html, Css technologies for easy handling and reading, being intuitive for the end user; the data collected for the operation of the system will be granted by the Ministry of Health. The implementation of this system has as its main objective to avoid the collapse of the health system and thus reduce the mortality rate in the population.

Descriptors: Data – Machine Learning – Heat Maps.

INDICE

APROBACION DEL ASESOR.....	III
DECLARACION DE AUTENTICIDAD.....	IV
APROBACION DEL TRIBUNAL.....	V
AGRADECIMIENTO	VI
DEDICATORIA	VII
RESUMEN.....	VIII
ABSTRACT.....	IX
SUMMARY	IX
CAPITULO I.....	1
INTRODUCCION	1
EL PROBLEMA	1
CONTEXTUALIZACION	3
ANTECEDENTES.....	5
CONTRIBUCIÓN TEÓRICA, ECONÓMICA, SOCIAL Y/O TECNOLÓGICA	7
Aplicación web.....	7
Procesos ETL	9
Depurar Información.....	11
Lenguaje de programación.....	11
Python	13
Django.....	14
Flask.....	15
Aprendizaje automático	15
Análisis predictivo	17
Algoritmo regresión polinomial.....	17
QGIS	19
Mapa de calor.....	19
JUSTIFICACION	21
OBJETIVOS	22
General.....	22

Específicos	22
CAPÍTULO II	23
METODOLOGÍA	23
DISEÑO DEL TRABAJO	23
ÁREA DE ESTUDIO	24
MODALIDAD DE INVESTIGACIÓN.....	24
TÉCNICAS E INSTRUMENTOS.....	25
ANÁLISIS E INTERPRETACIÓN DE RESULTADOS.....	25
CAPÍTULOS III.....	28
PROPUESTA Y RESULTADOS ESPERADOS	28
METODOLOGÍA DE DESARROLLO	28
Metodologías ágiles	28
Scrum	28
Kanban	29
Scrum vs Kanban	30
Selección de la Metodología de desarrollo	30
ANÁLISIS DEL SISTEMA ACTUAL	31
ESPECIFICACION DE REQUERIMIENTOS	31
PLAN DE DESARROLLO DE SOFTWARE.....	32
Organización del Proyecto	32
Historias de usuario.....	32
FASES METODOLOGÍA KANBAN	34
Fase 1 Kanban:.....	34
Fase 1, 2 Kanban:.....	35
Fase 3 Kanban:.....	37
Fase 4 Kanban:.....	38
Fase 5 Kanban:.....	39
ANÁLISIS DE RESTRICCIONES	41
ESTUDIO DE FACTIBILIDAD	41
Factibilidad Operativa.....	41
Factibilidad Técnica.....	42
Factibilidad Económica.....	42

ANÁLISIS ORIENTADO A OBJETOS	43
Diagramas de Casos de Uso	43
DISEÑO DE LA INTERFAZ DE USUARIO	43
Diseño de Entradas.....	43
CODIFICACIÓN	44
PUESTA EN MARCHA.....	45
RESULTADOS ESPERADOS	45
CAPITULO IV.....	46
CONCLUSIONES Y RECOMENDACIONES.....	46
CONCLUSIONES	46
RECOMENDACIONES	47
LECTURA CITADA	48
ANEXOS	50

ÍNDICE DE GRÁFICOS

Gráfico 1: Árbol de problemas.....	2
Gráfico 2 Representación de Scrum.....	28
Gráfico 3 Kanban	29
Gráfico 4 Scrum vs Kanban	30
Gráfico 5 Datos covid-19	35
Gráfico 6 Datos covid-19 procesados	36
Gráfico 7 Resultados de consola del algoritmo de predicción	37
Gráfico 8 Mapa de calor.....	38
Gráfico 9 Diseño aplicación web	40

ÍNDICE DE TABLAS

Tabla 1 Entrevista.	25
Tabla 2 Roles y Responsabilidades.....	32
Tabla 3 Fase 1 Kanban.....	34
Tabla 4 Fase 1, 2 Kanban.....	35
Tabla 5 Fase 1, 2, 3 Kanban.....	36
Tabla 6 Fase 1, 2, 3 Kanban.....	38
Tabla 7 Fase 2, 3 Kanban.....	39
Tabla 8 Fase 3 Kanban.....	40

CAPITULO I

INTRODUCCION

EL PROBLEMA

Actualmente, el honorable gobierno provincial de Tungurahua no cuenta con un sistema de automatizado de predicción, por ende, ninguno de sus departamentos está integrado en un sistema. El COE Provincial y funcionarios encargados de la toma de decisiones no pueden llevar el registro y control de enfermedades en este caso la pandemia del COVID-19.

Los cercos epidemiológicos se los realiza de una forma anticuada en la cual simplemente no se puede controlar las grietas que deja realizarlo de esa manera, debido al crecimiento de la pandemia del COVID-19 tiene la necesidad de desarrollar un sistema de predicción y cerco epidemiológico para mejorar la toma de decisiones y reducir los contagios para que el sistema de salud no colapse.

El tratamiento manual de la información se lo realiza en una hoja de cálculo Excel en la cual existen muchas inconsistencias en la información como consecuencia la falta de acciones contra una pandemia.

Árbol de problemas

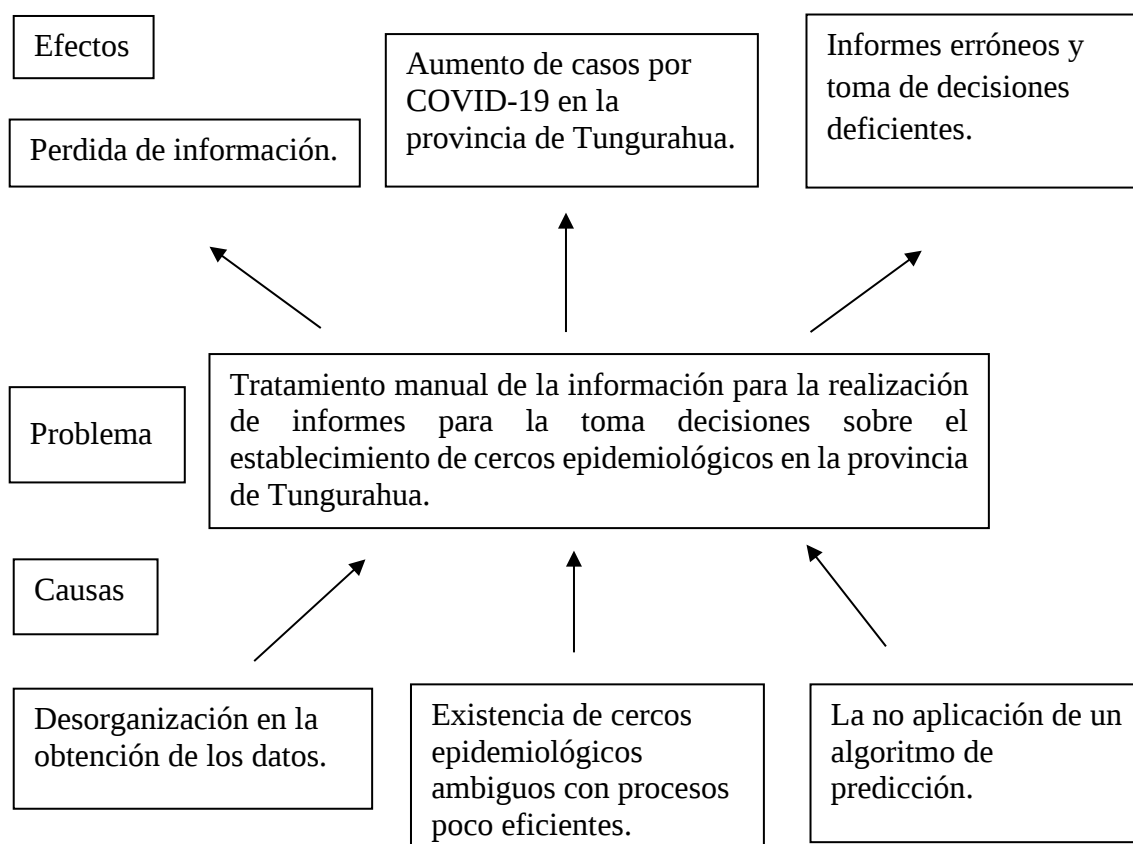


Gráfico 1: Árbol de problemas

Fuente: Investigador

Elaborado por: Kevin Estalin Llerena Ortiz

CONTEXTUALIZACION

Macro

La propagación de una enfermedad infecciosa es un problema de salud importante para los ciudadanos de todo el mundo. Identificar las fuentes de infección es clave para detener una pandemia. Es necesario actuar con rapidez y limitar las áreas geográficas de exposición a enfermedades. En noviembre y diciembre de 2019, la enfermedad altamente contagiosa conocida como enfermedad del coronavirus 2019 (COVID-19) se propagó por Wuhan, China. Desde entonces, se ha extendido a más de 6 millones de personas y 188 países de todo el mundo [1, Párr. 1]. Los coronavirus son una extensa familia de virus que pueden causar enfermedades tanto en animales como en humanos. En los seres humanos, se sabe que algunos coronavirus causan infecciones respiratorias que pueden variar desde el resfriado común hasta enfermedades más graves, como el síndrome respiratorio de Oriente Medio y el síndrome respiratorio agudo severo (SARS) [2, p. 1]. En el caso del COVID-19, su rápida expansión ha provocado muchas personas infectadas en muchos países debido a la falta de recursos sanitarios adecuados.

Meso

La pandemia de la COVID-19 [3, p. 1] según Esteban Ortiz-Prado y Raúl Fernández-Naranjo ha golpeado fuertemente al Ecuador, no solo en lo sanitario, sino en lo social, lo económico y lo político. Las estadísticas oficiales (que distan mucho de la realidad) son desalentadoras. Los entendidos sobre este tema, los familiares de los fallecidos y los propios afectados saben que este número dista mucho de la realidad. Con los análisis moleculares de acceso limitado, costosos y escasos es imposible hacer política pública. En vista de que la toma de decisiones requiere datos de alta calidad y ante la dificultad de acceder a los mismos, han aparecido múltiples fuentes de información y percepciones sobre el comportamiento de la pandemia en el país que no necesariamente se ajustan a los datos oficiales, y aún peor sin sistemas de predicción ni cercos epidemiológicos han

hecho que el sistema de salud colapse en poco tiempo.

Micro

En la provincia de Tungurahua [3, Párr. 5] la pandemia comenzó con la incertidumbre de no poder contener la propagación del virus con los manejos de registros manuales y la falta de sistemas que ayuden a la toma de decisiones, la importancia de acceder a información necesaria radica en que se confunde por mensajes opuestos. Así, muchas personas dejan de prestar atención a los informes. Esta (des)información puede acarrear falsa sensación de seguridad, prácticas inapropiadas y aumento de contagios en la provincia de Tungurahua.

ANTECEDENTES

Honorable Gobierno Provincial de Tungurahua

La Institución ha presentado un ciclo de apariciones y desapariciones dentro de los diferentes procesos constitucionales ejercidos entre los años de 1843 hasta la actualidad.

Los fondos con los que contaba el Gobierno Provincial correspondiente al 7% del valor de la venta de alcoholes de Tungurahua y tras el terremoto del 5 de agosto de 1949, este valor se incrementó al 20% para atender las necesidades más urgentes de los cantones. Con la recaudación de impuestos y la asignación económica que recibió del Estado, en el año de 1956 su presupuesto ascendió a los \$ 3753047,19, Entonces, el Gobierno Provincial presidido por el Sr. José Carrasco Mino, conocido como "Don Pepe", emprendió obras en beneficio de los tungurahueses como: varias escuelas, carreteras Inter parroquiales, y apoyo para la ejecución del Hospital de Píllaro y la Planta Eléctrica en el cantón de Baños.

Con el paso del tiempo, esta entidad va tomando importancia a medida que su presupuesto se va acrecentando con los ingresos que recibía de la Junta de Reconstrucción de Tungurahua, y desde entonces, su contribución para el desarrollo en aspectos de educación, electrificación, regadío, agua potable, vialidad, turismo, deporte, fondos dedicados para la construcción del Parque Industrial y otros servicios sociales ha permanecido notable hasta días presentes. El objetivo de instaurar gobiernos autónomos y descentralizados en cada capital de provincia es gestionar de una manera independiente, pública, política, social, económica y administrativa, en este caso, a la provincia de Tungurahua junto a sus nueve cantones. [4, Fig. 1]

MISION

Contribuir en la formación de ciudadanas y ciudadanos críticos con capacidad de análisis de la realidad, generación de espacios democráticos y con la discusión de propuestas que permitan alcanzar un desarrollo incluyente de la provincia.

VISION

Es una instancia del Honorable Gobierno Provincial de Tungurahua, que aporta en la construcción de ciudadanía, análisis de la realidad y la generación de espacios democráticos con la discusión de propuestas para el desarrollo de la Provincia y el país.

El departamento de tecnologías presidida por el Ing. Galo Robayo se desarrollan varios proyectos que en uno de ellos comprende la utilización del machine learning para predecir los contagios de COVID-19 y realizar cercos epidemiológicos para mejorar la toma de decisiones de los organismos encargados como el COE Provincial.

CONTRIBUCIÓN TEÓRICA, ECONÓMICA, SOCIAL Y/O TECNOLÓGICA

Aplicación web

Las aplicaciones web reciben este nombre porque se ejecutan en internet. Es decir que los datos o los archivos en los que trabajas son procesados y almacenados dentro de la web. Estas aplicaciones, por lo general, no necesitan ser instaladas en tu computador. [5, Párr. 1].

El concepto de aplicaciones web está relacionado con el almacenamiento en la nube. Toda la información se guarda de forma permanente en grandes servidores de internet y nos envían a nuestros dispositivos o equipos los datos que requerimos en ese momento, quedando una copia temporal dentro de nuestro equipo. [5, Párr. 2].

En cualquier momento, lugar y desde cualquier dispositivo podemos acceder a este servicio, sólo necesitamos una conexión a internet y nuestros datos de acceso, que por lo general son el nombre de usuario y contraseña. [5, Párr. 3].

Estos grandes servidores de internet que prestan el servicio de alojamiento están ubicados alrededor de todo el mundo, así hacen que el servicio prestado no sea tan costoso o gratuito en la mayoría de los casos y extremadamente seguro. [5, Párr. 4].

Una aplicación web (aplicación web) es un programa de aplicación que se almacena en un servidor remoto y se entrega a través de Internet a través de una interfaz de navegador. Los servicios web son aplicaciones web por definición y muchos sitios web, aunque no todos, contienen aplicaciones web. Según el editor de Web.AppStorm, Jarel Remick, cualquier componente del sitio web que realice alguna función para el usuario califica como una aplicación web. [5, Párr. 4]

Las aplicaciones web pueden diseñarse para una amplia variedad de usos y pueden ser utilizadas por cualquier persona; de una organización a un individuo por numerosas razones. Las aplicaciones web de uso común pueden incluir correo web, calculadoras en línea o tiendas de comercio electrónico. A algunas aplicaciones web solo se puede acceder mediante un navegador específico; sin embargo, la mayoría están disponibles sin importar el navegador. [5, Párr. 4]

Cómo funcionan las aplicaciones web

Las aplicaciones web no necesitan descargarse ya que se accede a ellas a través de una red. Los usuarios pueden acceder a una aplicación web a través de un navegador web como Google Chrome, Mozilla Firefox o Safari. [5, Párr. 6]

Para que una aplicación web funcione, necesita un servidor web, un servidor de aplicaciones y una base de datos. Los servidores web administran las solicitudes que provienen de un cliente, mientras que el servidor de aplicaciones completa la tarea solicitada. Se puede utilizar una base de datos para almacenar cualquier información necesaria. [5, Párr. 7]

Las aplicaciones web suelen tener ciclos de desarrollo cortos y se pueden crear con pequeños equipos de desarrollo. La mayoría de las aplicaciones web están escritas en JavaScript, HTML5 u hojas de estilo en cascada (CSS). La programación del lado del cliente generalmente utiliza estos lenguajes, que ayudan a construir un front-end de aplicaciones. La programación del lado del servidor se realiza para crear los scripts que utilizará una aplicación web. Los lenguajes como Python, Java y Ruby se usan comúnmente en la programación del lado del servidor. [5, Párr. 8]

Beneficios

Las aplicaciones web tienen muchos usos diferentes, y con esos usos vienen muchos beneficios potenciales. Algunos beneficios comunes de las aplicaciones web incluyen: [5, Párr. 9]

- Permitir que varios usuarios accedan a la misma versión de una aplicación.

- No es necesario instalar aplicaciones web.
- Se puede acceder a las aplicaciones web a través de varias plataformas, como una computadora de escritorio, una computadora portátil o un dispositivo móvil.
- Se puede acceder a través de múltiples navegadores.
- Aplicación web frente a otros tipos de aplicaciones [5, Párr. 9]

Dentro del sector de la informática móvil, las aplicaciones web a veces se contrastan con las aplicaciones nativas, que son aplicaciones que se desarrollan específicamente para una plataforma o dispositivo en particular y se instalan en ese dispositivo. Sin embargo, los dos no son mutuamente excluyentes. Las aplicaciones nativas son aplicaciones que normalmente se descargan y se crean específicamente para el tipo de dispositivo en el que se descargan. Las aplicaciones nativas normalmente pueden hacer uso del hardware específico del dispositivo, como un GPS o una cámara en una aplicación nativa móvil. [5, Párr. 10]

Los programas que combinan los dos enfoques a veces se denominan aplicaciones híbridas. Las aplicaciones híbridas funcionan de manera similar a una aplicación web, pero se instalan en el dispositivo como lo haría una aplicación nativa. Las aplicaciones híbridas también pueden aprovechar los recursos específicos del dispositivo mediante el uso de API internas. Las aplicaciones nativas descargadas a veces pueden funcionar sin conexión; sin embargo, las aplicaciones híbridas no tienen esta funcionalidad. Una aplicación híbrida normalmente compartirá elementos de navegación similares a los de una aplicación web, ya que se basan en aplicaciones web. [5, Párr. 11]

Procesos ETL

Los procesos ETL son una parte de la integración de datos, pero es un elemento importante cuya función completa el resultado de todo el desarrollo de la cohesión de aplicaciones y sistemas. La palabra ETL corresponde a las siglas en inglés de:

Extraer: extract. Transformar: transform. Y Cargar: load. Con ello, queremos decir que todo proceso ETL consta precisamente de estas tres fases: extracción, transformación y carga. [6, Párr. 1]. ETL, que significa extraer, transformar y cargar, es un proceso de integración de datos que combina datos de múltiples fuentes de datos en un único almacén de datos consistente que se carga en un almacén de datos u otro sistema de destino. A medida que las bases de datos crecieron en popularidad en la década de 1970, ETL se introdujo como un proceso para integrar y cargar datos para cálculo y análisis, convirtiéndose eventualmente en el método principal para procesar datos para proyectos de almacenamiento de datos. ETL proporciona la base para el análisis de datos y los flujos de trabajo de aprendizaje automático. A través de una serie de reglas comerciales, ETL limpia y organiza los datos de una manera que aborda necesidades específicas de inteligencia comercial, como informes mensuales, pero también puede abordar análisis más avanzados, que pueden mejorar los procesos de back-end o las experiencias del usuario final. ETL se utiliza a menudo por una organización para: Extraiga datos de sistemas heredados, Limpiar los datos para mejorar la calidad de los datos y establecer la coherencia, Cargar datos en una base de datos de destino [7, Párr. 2].

ETL es el proceso mediante el cual los datos se extraen de fuentes de datos (que no están optimizadas para análisis) y se trasladan a un host central (que sí lo está). Los pasos exactos de ese proceso pueden diferir de una herramienta ETL a otra, pero el resultado final es el mismo. [7, Párr. 3]

En su forma más básica, el proceso ETL abarca la extracción, transformación y carga de datos. Si bien la abreviatura implica un proceso ordenado de tres pasos: extraer, transformar, cargar, esta definición simple no captura: [7, Párr. 4]

- El transporte de datos
- La superposición entre cada una de estas etapas
- Cómo las nuevas tecnologías están cambiando este flujo [7, Párr. 4]

Los datos se extraen de bases de datos de procesamiento de transacciones en línea (OLTP), hoy más comúnmente conocidas como "bases de datos transaccionales", y otras fuentes de datos. Las aplicaciones OLTP tienen un alto rendimiento, con un gran número de solicitudes de lectura y escritura. No se prestan bien al análisis de datos o tareas de inteligencia empresarial. Luego, los datos se transforman en un área de ensayo. Estas transformaciones cubren tanto la limpieza de datos como la optimización de los datos para el análisis. Luego, los datos transformados se cargan en una base de datos de procesamiento analítico en línea (OLAP), hoy más comúnmente conocida como una base de datos de análisis. [7, Párr. 5]

Luego, los equipos de inteligencia comercial (BI) ejecutan consultas sobre esos datos, que finalmente se presentan a los usuarios finales o a las personas responsables de tomar decisiones comerciales, o se utilizan como entrada para algoritmos de aprendizaje automático u otros proyectos de ciencia de datos. Un problema común que se encuentra aquí es que, si los resúmenes de OLAP no pueden admitir el tipo de análisis que el equipo de BI quiere hacer, entonces todo el proceso debe ejecutarse nuevamente, esta vez con diferentes transformaciones. [7, Párr. 6]

Depurar Información

La información de depuración es una colección de información generada por el compilador o programa ensamblador de C / C ++ que describe la aplicación que debe utilizar un depurador. [8, Párr. 1].

Lenguaje de programación

Los lenguajes de programación consisten en código, es decir, un grupo de oraciones escritas para la computadora que lo dicen qué hacer. El código no es mágico; de hecho, parece un extraño hijo bastardo del inglés y la lógica. bastante a menudo. Sin embargo, a diferencia de los lenguajes naturales como el inglés o el chino, los lenguajes formales como Python se esfuerza por eliminar la ambigüedad. Tienen

que eliminar la ambigüedad, porque se supone que su código para decirle a la computadora exactamente qué hacer. No quieres que tu computadora haga nada improvisado decisiones por ti, ¿verdad? Como cualquier idioma extranjero, una vez que lees el código lo suficiente, cualquier fragmento de código encontrará que empezará a tener sentido, y si no es así, con un poco de investigación en Internet y una buena libro de referencia o dos, el código tendrá sentido. Aun así, la computadora no lee directamente su código. Tiene que traducirlo en algo aún más indescifrable para los simples mortales: el código máquina de ceros y unos, registrar comandos, etc. Para hacer esto necesita un intérprete, un programa que le escribe y ejecuta código (¡o un programa que hace programas!). Los intérpretes le permiten escribir y ejecutar código inmediatamente. Los compiladores le permiten convertir su código directamente en código de máquina para ejecutarlo más tarde. [9, p. 1]

Un lenguaje de programación es cualquier conjunto de reglas que convierte cadenas, o elementos de programa gráficos en el caso de lenguajes de programación visuales, en varios tipos de salida de código de máquina. Los lenguajes de programación son un tipo de lenguaje informático y se utilizan en la programación informática para implementar algoritmos. [9, p. 1]

La mayoría de los lenguajes de programación consisten en instrucciones para computadoras. Hay máquinas programables que utilizan un conjunto de instrucciones específicas, en lugar de lenguajes de programación generales. Desde principios del siglo XIX, se han utilizado programas para dirigir el comportamiento de máquinas como telares Jacquard, cajas de música y pianolas. [9, p. 2] Los programas para estas máquinas (como los pergaminos de una pianola) no produjeron un comportamiento diferente en respuesta a diferentes entradas o condiciones. [9, p. 2]

Se han creado miles de lenguajes de programación diferentes, y cada año se crean más. Muchos lenguajes de programación están escritos en forma imperativa (es decir, como una secuencia de operaciones a realizar) mientras que otros lenguajes usan la forma declarativa (es decir, se especifica el resultado deseado, no cómo

lograrlo). [9, p. 2]

La descripción de un lenguaje de programación generalmente se divide en los dos componentes de sintaxis (forma) y semántica (significado), que generalmente están definidos por un lenguaje formal. Algunos lenguajes están definidos por un documento de especificación (por ejemplo, el lenguaje de programación C está especificado por un estándar ISO), mientras que otros lenguajes (como Perl) tienen una implementación dominante que se trata como una referencia. Algunos lenguajes tienen ambos, siendo comunes el lenguaje básico definido por un estándar y las extensiones tomadas de la implementación dominante. [9, p. 2]

La teoría del lenguaje de programación es un subcampo de la informática que se ocupa del diseño, implementación, análisis, caracterización y clasificación de los lenguajes de programación. [9, p. 2]

Python

Python, un lenguaje de programación de muy alto nivel, orientado a objetos y de código abierto, diseñado para optimizar la velocidad de desarrollo. Aunque es completamente de propósito general, Python se llama a menudo en el lenguaje de programación orientado a objetos, en parte porque por su gran facilidad de uso, y en parte porque se utiliza comúnmente para orquestar o "pegar" otros componentes de software en una aplicación. [10, p. 1]

Python es un lenguaje de programación interpretado, orientado a objetos de alto nivel y con semántica dinámica. Su sintaxis hace énfasis en la legibilidad del código, lo que facilita su depuración y, por tanto, favorece la productividad. Ofrece la potencia y la flexibilidad de los lenguajes compilados con una curva de aprendizaje suave. Aunque Python fue creado como lenguaje de programación de uso general, cuenta con una serie de librerías y entornos de desarrollo para cada una de las fases del proceso de Data Science. Esto, sumado a su potencia, su carácter

open source y su facilidad de aprendizaje le ha llevado a tomar la delantera a otros lenguajes propios de la analítica de datos por medio de Machine Learning como pueden ser SAS (software comercial líder hasta el momento) y R (también open source, pero más propio de entornos académicos o de investigación). Python fue creado por Guido Van Rossum en 1991 y, como curiosidad, debe su nombre a la gran afición de su creador por las películas del grupo Monty Python. Además de librerías de herramientas científicas, numéricas, de herramientas de análisis y estructuras de datos, o de algoritmos de Machine Learning como NumPy, SciPy, Matplotlib, Pandas o PyBrain, Python ofrece entornos interactivos de programación orientados al Data Science. [11, p. 1]

Ventajas de elegir Python web

El código Python tiene numerosas ventajas a la hora de sentarse a crear webs. Te detallamos algunas de ellas a continuación: [11, p. 2]

- Es lenguaje open source, lo que quiere decir que es de software libre. Podrás usarlo gratuitamente en todo momento.
- Es multiplataforma, válido para todo tipo de sistemas y dispositivos.
- Es entendible para todos los programadores. La forma de escribir código Python es ordenada y limpia, por lo que cualquiera sabrá entender y trabajar sobre tu estructura creada.
- Tiene un estilo flexible. Con Python web podrás crear multitud de elementos, desde listas hasta funciones más complejas.
- Estilo sano de programación. Python web está pensado para que funcione siguiendo unas reglas concretas, de forma que todo el mundo siga las mismas directrices. [11, p. 2]

Django

Django es un marco web Python de alto nivel que fomenta un desarrollo rápido y un diseño limpio y pragmático. Creado por desarrolladores experimentados, se

ocupa de gran parte de las molestias del desarrollo web, por lo que puede concentrarse en escribir su aplicación sin necesidad de reinventar la rueda. Es gratis y de código abierto. [11, p. 4]

Características

- Ridículamente rápido: Django fue diseñado para ayudar a los desarrolladores a llevar las aplicaciones desde el concepto hasta su finalización lo más rápido posible. [11, p. 6]
- Tranquilizadoramente seguro: Django se toma la seguridad en serio y ayuda a los desarrolladores a evitar muchos errores de seguridad comunes. [11, p. 6]
- Extremadamente escalable: Algunos de los sitios más concurridos de la web aprovechan la capacidad de Django para escalar de forma rápida y flexible. [11, p. 6]

Flask

Flask es un marco web, es un módulo de Python que le permite desarrollar aplicaciones web fácilmente. Tiene un núcleo pequeño y fácil de ampliar: es un micro marco que no incluye un ORM (Object Relational Manager) o características similares. [11, p. 7]

Tiene muchas características interesantes como enrutamiento de URL, motor de plantillas. Es un marco de aplicación web WSGI. [11, p. 7]

Aprendizaje automático

El aprendizaje automático es el estudio de algoritmos informáticos que mejoran automáticamente a través de la experiencia y mediante el uso de datos. Se considera parte de la inteligencia artificial. Los algoritmos de aprendizaje automático

construyen un modelo basado en datos de muestra, conocidos como "datos de entrenamiento", con el fin de hacer predicciones o decisiones sin estar programados explícitamente para hacerlo. Los algoritmos de aprendizaje automático se utilizan en una amplia variedad de aplicaciones, como en medicina, filtrado de correo electrónico, reconocimiento de voz y visión por computadora, donde es difícil o inviable desarrollar algoritmos convencionales para realizar las tareas necesarias. Un subconjunto del aprendizaje automático está estrechamente relacionado con las estadísticas computacionales, que se enfoca en hacer predicciones usando computadoras; pero no todo el aprendizaje automático es aprendizaje estadístico. El estudio de la optimización matemática aporta métodos, teoría y dominios de aplicación al campo del aprendizaje automático. La minería de datos es un campo de estudio relacionado, que se centra en el análisis de datos exploratorios a través del aprendizaje no supervisado. En su aplicación a los problemas empresariales, el aprendizaje automático también se conoce como análisis predictivo. [12]

El aprendizaje automático es una rama de la inteligencia artificial (IA) y la informática que se centra en el uso de datos y algoritmos para imitar la forma en que aprenden los humanos, mejorando gradualmente su precisión. [12]

IBM tiene una rica historia con el aprendizaje automático. A uno de los suyos, Arthur Samuel, se le atribuye haber acuñado el término "aprendizaje automático" con su investigación (PDF, 481 KB) (el enlace se encuentra fuera de IBM) en torno al juego de damas. Robert Nealey, el autoproclamado maestro de las damas, jugó en una computadora IBM 7094 en 1962 y perdió contra la computadora. En comparación con lo que se puede hacer hoy, esta hazaña parece casi trivial, pero se considera un hito importante en el campo de la inteligencia artificial. Durante las próximas dos décadas, los desarrollos tecnológicos en torno al almacenamiento y la potencia de procesamiento permitirán algunos de los productos innovadores que conocemos y amamos hoy, como el motor de recomendación de Netflix o los autos sin conductor. [12]

El aprendizaje automático es un componente importante del creciente campo de la

ciencia de datos. Mediante el uso de métodos estadísticos, se entrenan algoritmos para realizar clasificaciones o predicciones, descubriendo información clave dentro de proyectos de minería de datos. Estos conocimientos luego impulsan la toma de decisiones dentro de las aplicaciones y los negocios, lo que idealmente impacta en las métricas de crecimiento clave. A medida que el big data continúe expandiéndose y creciendo, aumentará la demanda del mercado de científicos de datos, lo que requerirá que ayuden a identificar las preguntas comerciales más relevantes y, posteriormente, los datos para responderlas. [12]

Análisis predictivo

El análisis predictivo es un término paraguas para referirnos al conjunto de procesos que implican aplicar diferentes técnicas computacionales con el objetivo de realizar predicciones sobre el futuro basándonos en datos pasados. La variedad de técnicas empleadas incluye minería de datos (data mining), modelado, reconocimiento de patrones, graph analytics. [13, Párr. 4]

Algoritmo regresión polinomial

La regresión polinómica es un algoritmo de aprendizaje de regresión supervisado. El algoritmo de regresión establece un modelo de regresión entre variables y obtiene la correlación entre variables y variables dependientes a través del proceso de aprendizaje (entrenamiento). El análisis de regresión puede usarse para modelos predictivos o de clasificación. Los algoritmos de regresión comunes incluyen: Regresión lineal, Regresión no lineal, Regresión logística, Regresión polinómica, Regresión de cresta, Regresión de lazo y la regresión de ElasticNet. Entre ellos, la regresión lineal, la regresión no lineal y la regresión logística son las más utilizadas. En muchos casos, el modelo lineal no puede ajustarse bien a la curva de datos objetivo, lo que requiere la introducción de un modelo de regresión no lineal. Existen múltiples estrategias para la regresión no lineal. La primera estrategia es transformar la regresión no lineal en una regresión lineal; la segunda estrategia es

transformar la regresión no lineal en una regresión polinómica. En la regresión polinómica, se agrega la mayor potencia de la entidad (como el término cuadrado o cúbico), lo que equivale a aumentar el grado de libertad del modelo para capturar cambios no lineales en los datos. [14, Párr. 1,2]

In simple linear regression algorithm only works when the relationship between the data is linear but suppose if we have non-linear data then Linear regression will not capable to draw a best-fit line and it fails in such conditions. consider the below diagram which has a non-linear relationship and you can see the Linear regression results on it, which does not perform well means which does not come close to reality. Hence, we introduce polynomial regression to overcome this problem, which helps identify the curvilinear relationship between independent and dependent variables. [14, Párr. 1,2]

How Polynomial Regression Overcomes the problem of Non-Linear data?

Polynomial regression is a form of Linear regression where only due to the Non-linear relationship between dependent and independent variables we add some polynomial terms to linear regression to convert it into Polynomial regression. [14, Párr. 1,2]

Suppose we have X as independent data and Y as dependent data. Before feeding data to a model in preprocessing stage we convert the input variables into polynomial terms using some degree. [14, Párr. 1,2]

Consider an example my input value is 35 and the degree of a polynomial is 2 so I will find 35 power 0, 35 power 1, and 35 power 2 And this helps to interpret the non-linear relationship in data. [14, Párr. 1,2]

The degree of order which to use is a Hyperparameter, and we need to choose it wisely. But using a high degree of polynomial tries to overfit the data and for smaller values of degree, the model tries to underfit so we need to find the optimum value of a degree. [14, Párr. 1,2]

QGIS

QGIS (anteriormente llamado también Quantum GIS) es un Sistema de Información Geográfica (SIG) de software libre y de código abierto para plataformas GNU/Linux, Unix, Mac OS, Microsoft Windows y Android. Fue uno de los primeros ocho proyectos de la Fundación OSGeo y en 2008 oficialmente graduó de la fase de incubación. Permite manejar formatos raster y vectoriales a través de la biblioteca GDAL (GADL/OGR), así como bases de datos. [15, Párr. 1]

Mapa de calor

Un mapa de calor es una técnica de visualización de datos que muestra la magnitud de un fenómeno como color en dos dimensiones. La variación en el color puede ser por matiz o intensidad, dando pistas visuales obvias al lector sobre cómo el fenómeno se agrupa o varía en el espacio. Hay dos categorías fundamentalmente diferentes de mapas de calor: el mapa de calor del clúster y el mapa de calor espacial. En un mapa de calor de conglomerados, las magnitudes se presentan en una matriz de tamaño de celda fijo cuyas filas y columnas son fenómenos y categorías discretos, y la clasificación de filas y columnas es intencional y algo arbitraria, con el objetivo de sugerir conglomerados o representarlos como descubierto mediante análisis estadístico. El tamaño de la celda es arbitrario pero lo suficientemente grande como para ser claramente visible. Por el contrario, la posición de una magnitud en un mapa de calor espacial está forzada por la ubicación de la magnitud en ese espacio, y no existe la noción de células; se considera que el fenómeno varía continuamente. [16, Párr. 1]

Existen diferentes tipos de mapas de calor:

- Los mapas de calor de biología se utilizan normalmente en biología

molecular para representar el nivel de expresión de muchos genes en una serie de muestras comparables (por ejemplo, células en diferentes estados, muestras de diferentes pacientes) a medida que se obtienen de micro matrices de ADN. [16, Párr. 2]

- El mapa de árbol es una partición jerárquica 2D de datos que se parece visualmente a un mapa de calor. [16, Párr. 2]
- Un gráfico de mosaico es un mapa de calor en mosaico para representar una tabla de datos bidireccional o superior. Al igual que con los mapas de árbol, las regiones rectangulares en un gráfico de mosaico están organizadas jerárquicamente. Eso significa que las regiones son rectángulos en lugar de cuadrados. Friendly (1994) examina la historia y el uso de este gráfico. [16, Párr. 2]
- Una visualización de función de densidad es un mapa de calor para representar la densidad de puntos en un mapa. Permite percibir la densidad de puntos independientemente del factor de zoom. Perrot et al. (2015) propusieron una forma de usar la función de densidad para visualizar miles y miles de millones de puntos usando infraestructura de big data con Spark y Hadoop. [16, Párr. 2]

JUSTIFICACION

La importancia de automatizar procesos y utilizar la tecnología a favor de una población es la prioridad de los gobernantes, reduciendo riesgos en la salud de las personas y lo más importante ayudar en la toma de decisiones.

Es por ello que el presente proyecto es de gran interés para el gobierno provincial de Tungurahua, en el departamento de tecnologías ya que la misma tiene la necesidad de dar solución al inminente crecimiento de casos por COVID-19 en la provincia de Tungurahua mediante la creación de cercos epidemiológicos inteligentes y un sistema de predicción de contagios que facilitaran a las autoridades competentes la información necesaria para delimitar los sitios seguros y evitar así las zonas con mayores riesgos de contagio, ya que actualmente se utilizan simples hojas de Excel sin ninguna clase de orden lo que dificulta la lectura de datos así como la integridad de los mismos por ende la toma de decisiones no serán las mejores.

Ante la crisis sanitaria, económica y psicológica, los sectores más vulnerables son las pequeñas, medianas empresas y, las personas de mayor riesgo que son los adultos mayores, es por ello la necesidad de prevenir futuras pandemias y delimitar áreas de contagio. El problema surgió por la pandemia que no solo es en Ecuador si no global como lo indica la OMS hasta 120 millones de personas están en riesgo de enfermedad y muerte debido a los efectos secundarios de la pandemia de COVID-19. El informe de réplicas de Visión Mundial considera lo que sucedería si los devastadores impactos secundarios del brote de Ébola 2015-2016 en los niños se replicaran en los 24 países más frágiles cubiertos por el llamamiento humanitario COVID-19 de la ONU, y los problemas en Ecuador, así como en Tungurahua no solo son de carácter de salud si no también sociales entre otros impactos muy grandes en la economía.

Además, se innovará la manera en que los datos han sido tratados por administraciones anteriores, lo que otorgara mayor confianza en los ciudadanos al poseer nuevas herramientas que no solo funcionaran con esta pandemia sino con futuros problemas de salud mundial.

OBJETIVOS

General

- Desarrollar un sistema de predicción de contagios y cerco epidemiológico del Covid-19 en la provincia de Tungurahua para el Honorable Gobierno Provincial de Tungurahua.

Específicos

- Realizar un ETL (Extraer, Transformar y Cargar) a los datos de casos COVID-19 para depurar la información y poderlos utilizar en el algoritmo de minería de datos.
- Aplicar el algoritmo de regresión polinomial para predecir contagios de COVID-19 en la provincia de Tungurahua.
- Crear un mapa de calor de los casos COVID-19 con sus respectivas ubicaciones geográficas en una página web para realizar un cerco epidemiológico.

CAPÍTULO II

METODOLOGÍA

DISEÑO DEL TRABAJO

Debido a la pandemia del COVID-19 el honorable Gobierno provincial de Tungurahua se ve en la necesidad de implementar un sistema que ayude en los procesos internos para que sus autoridades puedan tener reportes de los cuales se pueda mejorar la toma de decisiones ayudando al sistema de salud y a la ciudadanía en general en controlar el aumento de contagios en la provincia. Claramente el mundo el país y la provincia de Tungurahua no estaban preparados para tal pandemia; la recolección de datos carece totalmente de organización y transparencia. El centro de operaciones de emergencia o COE provincial al no tener claros los datos de contagios junto con reportes de los mismos no pueden tomar las decisiones correctas que necesita la provincia de Tungurahua para detener la curva de contagios de COVID-19. A los datos se les efectuará un ETL (extracción, transformación y lectura) para organizar la información en un archivo CSV (valores separados por comas) en Excel a la información proporcionada por el departamento de tecnologías del Gobierno provincial de Tungurahua con lo cual los datos estarán listos para poder ser utilizados por el sistema integral. El sistema que se desarrollará tiene un diseño basado en estas pautas primeramente se aplicará un algoritmo de predicción llamado algoritmo de regresión polinomial que ayudará a predecir los contagios en la provincia de Tungurahua para controlar la curva de contagios y mantener los datos reales por debajo de esta predicción para así poder ayudar al sistema de salud a no colapsar; este algoritmo se desarrollará en el lenguaje de programación Python que es la última tendencia y tecnología para el uso de la inteligencia artificial también contará con un mapa de calor el cual ayudará a establecer cercos epidemiológicos en la provincia de Tungurahua limitando las zonas de mayor probabilidad de contagios; esta herramienta se la desarrollará en QGIS y estos dos módulos estarán integradas en una aplicación web igualmente desarrollada con Django como back end un framework que utiliza Python para la creación de aplicaciones web así como VUE para el front end en la cual autoridades

podrán generar reportes y predecir los contagios de COVID-19, la ciudadanía también podrá hacer uso de este sistema si así lo deciden las autoridades.

ÁREA DE ESTUDIO

El área de estudio presente en el trabajo de titulación es:

Ingeniería de Software en la cual hace presencia la creación de software confiable y de calidad, proporcionando soporte operacional y de mantenimiento. Bases de datos aplicando un ETL (extracción, transformación y lectura). Inteligencia artificial utilizando algoritmos de predicción. Minería de Datos intentando descubrir patrones de contagios. Programación Avanzada desarrollando una aplicación web con Django y Vuejs.

MODALIDAD DE INVESTIGACIÓN

El proyecto es de modalidad especial por la inesperada pandemia de COVID-19 que golpeo fuerte en la provincia de Tungurahua, es por ello la necesidad de disminuir la curva de contagios y ayudar al sistema de salud y a las entidades de emergencia en la toma de decisiones.

Bibliográfica o documental, puesto que se realizó la consulta de libros, web para fundamentar teóricamente bases conceptuales relacionadas con el proyecto, apoyándose en fuentes de carácter documental esto es, en documentos de cualquier especie tales como, las obtenidas a través de fuentes bibliográficas basada en la consulta de libros y hemerográficas basada en artículos o ensayos de revistas.

Para recolectar los datos de contagios se utilizó la modalidad de campo, esta información fue proporcionada por el departamento de tecnologías del honorable gobierno provincial de Tungurahua; a su vez se utilizó la modalidad bibliográfica de donde se extrajo información valiosa con la cual se desarrolló el algoritmo de predicción.

TÉCNICAS E INSTRUMENTOS

Se aplica la técnica de la entrevista para el levantamiento de la información, así como requerimientos del sistema y posibilidades en cuanto a la disponibilidad de datos.

Entrevistador: Kevin Llerena

Entrevistado: Ing. Galo Robayo

La entrevista fue de tipo semi estructurada por la aplicación de Zoom.

ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

Luego de aplicar la entrevista al Ingeniero encargado de la información del departamento de tecnologías del gobierno provincial de Tungurahua se muestra un resumen del resultado de la misma en la tabla 1 que se presenta a continuación.

Tabla 1 Entrevista.

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

Pregunta:	Respuesta:	Interpretación:
#1: ¿Cuáles cree usted que son las necesidades que tiene la prefectura y las autoridades de emergencia para enfrentar la pandemia del covid-19?	Yo creo que las autoridades del COE Provincial ante tal situación que atraviesa el país necesitan la ayuda de la tecnología para sobrellevar esta emergencia sanitaria y así mejorar la toma de decisiones con ayuda de un sistema que prediga los contagios y realice cercos epidemiológicos.	El jefe del departamento de tecnologías siente que las autoridades de emergencia necesitan manejar los datos con nuevas tecnologías que ayuden en la toma de decisiones.

#2: ¿Como y en que formato recogen los datos de contagios de covid-19 en la provincia o el ministerio de salud les proporciona estos datos para que las autoridades puedan tomar decisiones?	Los datos son confidenciales y nos proporciona el ministerio de salud, están guardados en un Excel y solo contienen las direcciones de los contagiados.	Los datos los manejan de forma manual y utilizan una hoja de cálculo en la cual no hay latitud ni longitud que se necesitarían para realizar los cercos epidemiológicos con mapas digitales.
#3: ¿Existe alguna manera en la que utilicen el machine learning para predecir futuros contagios de covid-19 en la provincia?	En el departamento de tecnologías del consejo provincial no existe algo similar y simplemente no se puede generar reportes para las autoridades todo se analiza manualmente.	Los datos los analizan manualmente por ende no existe un patrón a seguir por parte de las autoridades lo que afectaría de forma negativa en la toma de decisiones.
#4: ¿Qué tecnologías me recomendaría utilizar para el desarrollo del sistema de predicción de contagios y cercos epidemiológicos?	La verdad te recomendaría utilizar Matlab para realizar las predicciones con algún algoritmo de machine learning y para los mapas de cercos epidemiológicos utilizar ArcGIS y eso mostrar en una interfaz web.	Se deberá analizar con que tecnologías trabajar para el desarrollo del sistema priorizando software libre.
#5: ¿La información brindada por el sistema se compartirá con la población?	En cierto punto sí, pero eso depende de las políticas de las autoridades y la	La necesidad de mostrar tal información en la web era una inquietud para analizar en que se desarrollara y como se

	confidencialidad para manejos internos.	mostrarán las predicciones y cercos epidemiológicos.
#6: ¿Se necesitará la generación de reportes?	Claro que sí, la generación de reportes es muy necesario para que las autoridades puedan tomar decisiones basadas en las predicciones otorgadas por el sistema.	La generación de reportes será algo indispensable en el sistema a desarrollar.

CAPÍTULOS III

PROPUESTA Y RESULTADOS ESPERADOS

METODOLOGÍA DE DESARROLLO

Metodologías ágiles

Scrum

Fue desarrollado por Jeff Sutherland y Ken Schwaber a principios de los 90. En términos simples, Scrum descompone las organizaciones en pequeños equipos auto-organizados. Estos equipos rompen las funciones que han sido asignadas a entregar en pequeñas tareas manejables que abordan en iteraciones en tiempo real llamadas sprints. El más prescriptivo de los tres marcos considerados en este post, Scrum insiste en que los equipos comprenden tres roles distintos: el Product Owner, el ScrumMaster y los miembros del equipo de desarrollo. También hay cuatro reuniones prescritas en Scrum: la Reunión de Planificación de Sprint, el Daily Stand Up, la Sprint Review y la Sprint Retrospective. [17].

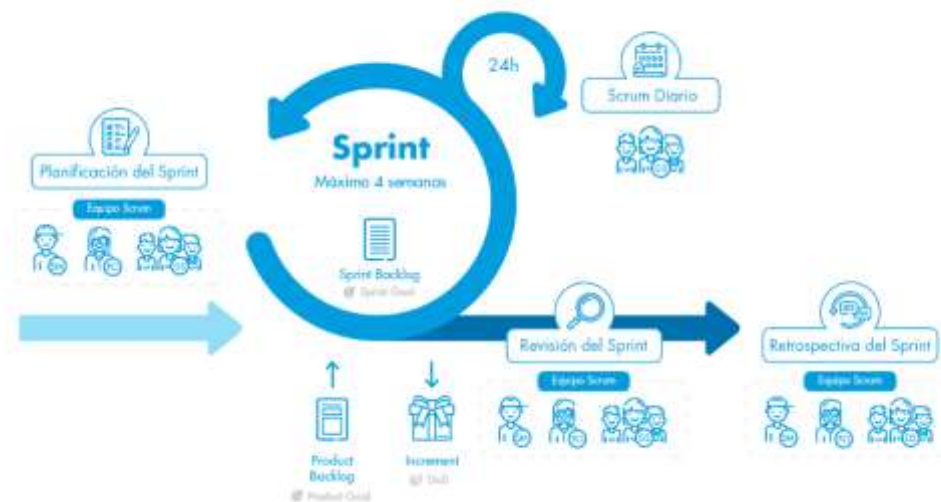


Gráfico 2 Representación de Scrum

Fuente: <https://netmind.net/es/scrum-el-pasado-y-el-futuro/>.

Elaborado por: Kevin Estalin Llerena Ortiz.

En el gráfico 1 se puede visualizar el ciclo de la metodología scrum junto con los autores que dicha metodología necesita para funcionar añadiendo algunas novedades y características asemejándose a Kanban, pero nunca serán lo mismo.

Kanban

Kanban es un sistema de producción tan eficiente como efectivo. Forma parte de las metodologías ágiles y su objetivo es gestionar la realización de las tareas hasta su finalización. [19, Párr. 1]

Kanban es una palabra japonesa formada por Kan, que quiere decir visual, y Ban, que significa tarjeta. Por lo tanto, Kanban hace referencia a las tarjetas visuales. [19, Párr. 2]

Esta metodología es muy sencilla, se puede actualizar y los equipos de trabajo la pueden asumir sin problema. Al ser un método visual permite que con un simple vistazo se conozca el estado de los proyectos y se puedan asignar nuevas tareas de manera muy efectiva. Para aplicarlo, es necesario un tablero de tareas con el que poder mejorar el trabajo y tener un ritmo sostenible. [19, Párr. 3]



Gráfico 3 Kanban

Fuente: <https://tecnosoluciones.com/10-razones-para-usar-la-metodologia-kanban-en-tu-organizacion/>.

Elaborado por: Kevin Estalin Llerena Ortiz.

En el gráfico 2 podemos observar la metodología Kanban lo simple y efectiva que es con sus 3 fases, Kanban ha sido descrita como un enfoque ágil para la gestión de proyectos. Principalmente se basa en el flujo continuo de trabajo cuyo objetivo es

gestionar, de manera general, cómo se van completando las tareas de un proyecto determinado [20, Párr. 2]

Scrum vs Kanban

SCRUM vs KANBAN		
	SCRUM	KANBAN
TIPO DE TRABAJO	Proyectos	Servicios y proyectos
EQUIPO	Interfuncional y autoorganizado	No requiere interfuncionalidad y autoorganización
ROLES EXIGIDOS	Sí	No
PERIODICIDAD DE LOS EVENTOS	Marcada por los sprints	Según se necesite
PETICIONES	Pueden esperar hasta el próximo sprint	Pueden haber peticiones muy urgentes

www.itnove.com

Gráfico 4 Scrum vs Kanban

Fuente: <https://itnove.com/>.

Elaborado por: Kevin Estalin Llerena Ortiz.

En el grafico 3 se puede visualizar una comparación entre las dos metodologías que se presenta en esta investigación; y por la cual Kanban fue seleccionado con sus características más fuertes que son: sus roles no exigidos y la periodicidad de los eventos según se necesite, con esta metodología ágil podemos realizar un proyecto de forma efectiva, asertiva y en el menor tiempo posible.

Selección de la Metodología de desarrollo

Kanban

La diferencia es que Scrum posee características que no necesita este proyecto y al hacer hincapié en lo ágil y eficiente que tiene que ser un proyecto se ha escogido Kanban porque es la metodología que más se adapta al proyecto ir desarrollando el

proyecto según vayan apareciendo nuevas necesidades y poderlas trabajar con urgencia sin exigir roles al no requerir Inter funcionalidad y autoorganización.

ANÁLISIS DEL SISTEMA ACTUAL

Debido a la pandemia que atraviesa el mundo muchos países han sido muy afectados, en Ecuador y específicamente en la provincia de Tungurahua los procesos que se realizan para tratar de frenar la curva de contagios se los hace de formas no automatizadas, utilizan una hoja de cálculo Excel en la cual la información no estaba preparada para realizar reportes automatizados ya que al estar completamente desorganizados dichos datos conlleva a malas interpretaciones y toma de decisiones erróneas; sin el correcto manejo de los datos es imposible realizar cercos epidemiológicos y predicciones; Es por ello que se pretende automatizar estos procesos ayudando a mejorar la toma de datos organizándolos, transformándolos para que su lectura ayude de mejor manera a la generación de reportes, aplicar los algoritmos de predicción y poder crear cercos epidemiológicos con mapas de calor geográficos los cuales ayudaran en la toma de decisiones por parte de las autoridades.

ESPECIFICACION DE REQUERIMIENTOS

- Transformar los datos en coordenadas geográficas y guardarlo en un archivo CSV para realizar el mapa de cercos epidemiológicos.
- Transformar los datos en número de casos por día y guardarlo en un archivo CSV para poder aplicar el algoritmo de predicción.
- Actualizar los puntos de latitud y longitud en el mapa de cercos epidemiológicos.

PLAN DE DESARROLLO DE SOFTWARE

Organización del Proyecto

Roles y Responsabilidades

Tabla 2 Roles y Responsabilidades

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

Cargo	Funciones
Jefe de Proyecto	Ing. Galo Robayo
Analista de Sistemas	Ing. Franklin Castillo
Programador	Kevin Llerena

Historias de usuario

Historia de usuario	
Número: 1	Usuario: Jefe de proyecto
Nombre de historia: Proceso ETL	
Prioridad en negocio: Alta	Riesgo en desarrollo: Alta
Puntos estimados: 2	Iteración asignada: 1
Programador responsable: Kevin Llerena	
Descripción: Se deberá analizar los datos entregados del Covid-19 por parte del Gobierno provincial de Tungurahua y aplicar un proceso ETL.	
Validación: Los datos estarán descritos de forma clara y concisa aplicada ya la depuración de los datos.	

Historia de usuario	
Número: 2	Usuario: Jefe de proyecto
Nombre de historia: Organizar datos csv	
Prioridad en negocio: Alta	Riesgo en desarrollo: Alta
Puntos estimados: 2	Iteración asignada: 1
Programador responsable: Kevin Llerena	
Descripción: Organizar los datos en un archivo csv para poder ser utilizados en un algoritmo de predicción.	
Validación: Archivo csv organizado y separado por comas.	

Historia de usuario	
Número: 3	Usuario: Jefe de proyecto
Nombre de historia: Algoritmo de predicción	
Prioridad en negocio: Alta	Riesgo en desarrollo: Alta
Puntos estimados: 2	Iteración asignada: 1
Programador responsable: Kevin Llerena	
Descripción: Como jefe de proyecto se requiere programar en Python el algoritmo de predicción.	
Validación: El algoritmo debe predecir los contagios a 2, 7 y 15 días	

Historia de usuario	
Número: 4	Usuario: Jefe de proyecto
Nombre de historia: Cerco epidemiológico	
Prioridad en negocio: Alta	Riesgo en desarrollo: Alta
Puntos estimados: 2	Iteración asignada: 1
Programador responsable: Kevin Llerena	
Descripción: Se requiere crear un mapa de calor con QGIS para los cercos epidemiológicos.	
Validación: Se debe poder observar las zonas con más contagios en Tungurahua de color rojo.	

Historia de usuario	
Número: 5	Usuario: Jefe de proyecto
Nombre de historia: Interfaz web	
Prioridad en negocio: Alta	Riesgo en desarrollo: Alta
Puntos estimados: 2	Iteración asignada: 1
Programador responsable: Kevin Llerena	
Descripción: Se requiere desarrollar una aplicación web con Django implementando el algoritmo de predicción.	
Validación: Se deberá poder observar en un sitio web la predicción realizada junto con el mapa de calor.	

FASES METODOLOGÍA KANBAN

Tabla 3 Fase 1 Kanban

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

Fecha: 13/06/2020		
Que hacer	Haciendo	Hecho
Historia de usuario: • 1 • 2 • 3 • 4 • 5		

Fase 1 Kanban:

En esta fase o etapa llamada “Que hacer” visualizamos nuestras historias de usuario en espera de poder ser ejecutadas por el investigador para seguir avanzando en proyecto de una forma ordenada y ágil.

Tabla 4 Fase 1, 2 Kanban

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

Fecha: 14/06/2020		
Que hacer	Haciendo	Hecho
Historia de usuario: • 3 • 4 • 5	Historia de usuario: • 1 • 2	

Fase 1, 2 Kanban:

En esta etapa podemos observar que la fase 2 “Haciendo” ya tenemos las dos primeras tareas las cuales están siendo ejecutadas por el investigador, a la vez tenemos 3 tareas pendientes en “Que hacer”.

Historia de usuario 1:

Los datos son receptados y se procede a analizar dichos datos encontrando muchas incongruencias y redundancia de datos como se puede observar en la siguiente imagen.

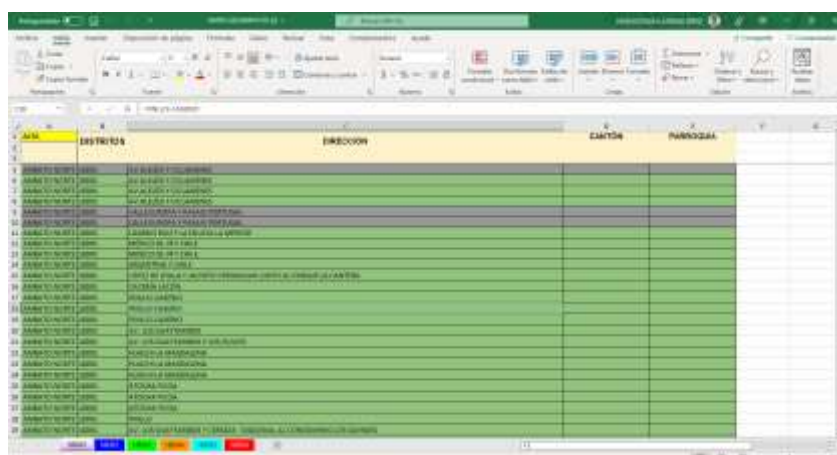


Gráfico 5 Datos covid-19

Fuente: Ministerio de salud – Ecuador.

Elaborado por: Kevin Estalin Llerena Ortiz.

En el gráfico 5 se observa los datos redundantes sin ninguna descripción del lugar y estado por ende se necesita de un proceso ETL.

Historia de usuario 2:

Los datos luego del proceso ETL se procede a organizar los datos en un archivo csv para su posterior carga en el algoritmo de predicción.

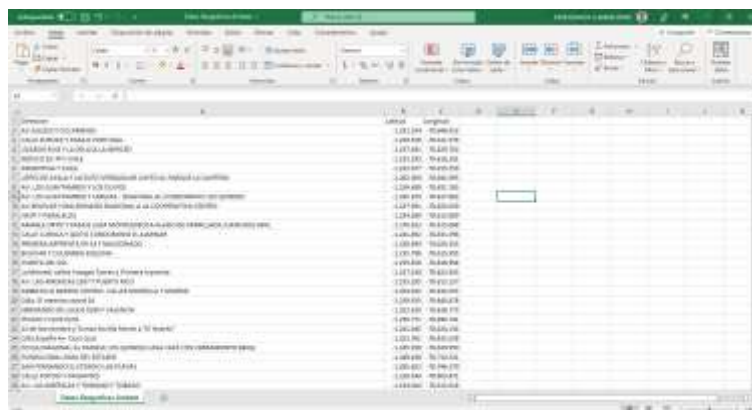


Gráfico 6 Datos covid-19 procesados

Fuente: Ministerio de salud pública.

Elaborado por: Kevin Estalin Llerena Ortiz.

En el grafico 6 se observan los datos ya organizados y listos para ser utilizados en el sistema junto con sus direcciones geográficas para posteriormente poder ser utilizadas en los mapas de calor.

Tabla 5 Fase 1, 2, 3 Kanban

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

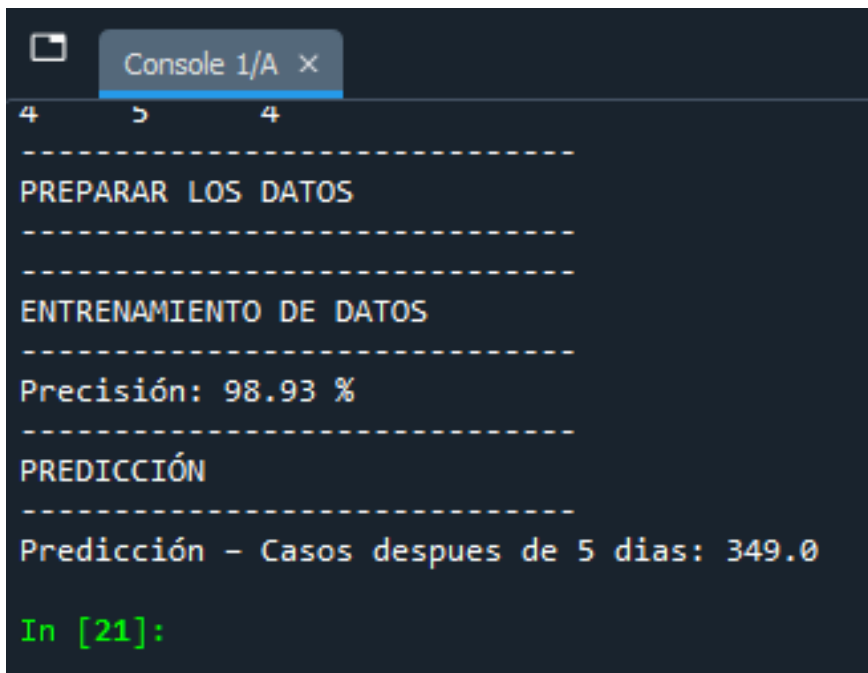
Fecha: 07/07/2021		
Que hacer	Haciendo	Hecho
Historia de usuario: • 4 • 5	Historia de usuario: • 3	Historia de usuario: • 1 • 2

Fase 3 Kanban:

En esta etapa podemos observar que la fase 3 “Haciendo” ya tenemos solo una tarea1 la cual está siendo ejecutada por el investigador, a la vez tenemos 2 tareas pendientes en “Que hacer”.

Historia de usuario 3:

Una vez finalizado el proceso de ETL y organización en un archivo csv se procede a la programación en Python de un algoritmo de predicción en el cual se utiliza un algoritmo de regresión polinomial siendo este uno de los más sencillos y eficaz dando como resultado una precisión del 98,93%.



```
Console 1/A x
4 5 4
-----
PREPARAR LOS DATOS
-----
ENTRENAMIENTO DE DATOS
-----
Precisión: 98.93 %
-----
PREDICCIÓN
-----
Predicción - Casos despues de 5 dias: 349.0
In [21]:
```

Gráfico 7 Resultados de consola del algoritmo de predicción

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

En el grafico 7 podemos observar los resultados de consola del algoritmo de predicción luego de ser entrenado con los datos que hemos organizado anteriormente, como se observa el algoritmo obtiene una precisión del 98,93% así como la predicción de casos covid-19 luego de 5 días.

Tabla 6 Fase 1, 2, 3 Kanban

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

Fecha: 07/08/2021		
Que hacer	Haciendo	Hecho
Historia de usuario: • 5	Historia de usuario: • 4	Historia de usuario: • 1 • 2 • 3

Fase 4 Kanban:

En esta etapa podemos observar que la fase 4 “Haciendo” ya tenemos solo una tarea la cual está siendo ejecutada por el investigador, a la vez tenemos 1 tarea pendiente en “Que hacer”.

Historia de usuario 4:

Con el mismo archivo csv que anteriormente se procesó lo utilizaremos en QGIS para realizar un mapa de calor con el cual se podrá establecer zonas seguras y formar un cerco epidemiológico.

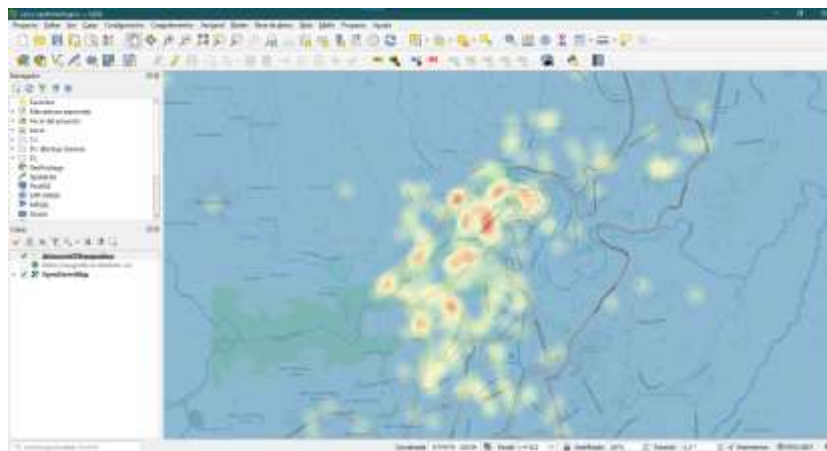


Gráfico 8 Mapa de calor

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz

En el grafico 8 podemos observar como se creo el mapa de calor para luego ser utilizado como cercos epidemiológicos elaborado en QGIS con el mapa geográfico de Tungurahua y los puntos geográficos de los casos Covid-19.

Tabla 7 Fase 2, 3 Kanban

Fuente: El investigador.

Elaborado por: Kevin Estalin Llerena Ortiz.

Fecha: 25/08/2021		
Que hacer	Haciendo	Hecho
	Historia de usuario: • 5	Historia de usuario: • 1 • 2 • 3 • 4

Fase 5 Kanban:

En esta etapa podemos observar que la fase 5 “Haciendo” ya tenemos solo una tarea la cual está siendo ejecutada por el investigador, a la vez ya no tenemos tareas pendientes en “Que hacer”.

Historia de usuario 5:

Luego de obtener los implementos necesarios se pone en marcha el desarrollo de una aplicación web en la cual se podrá visualizar de una manera intuitiva lo ya implementado anteriormente.

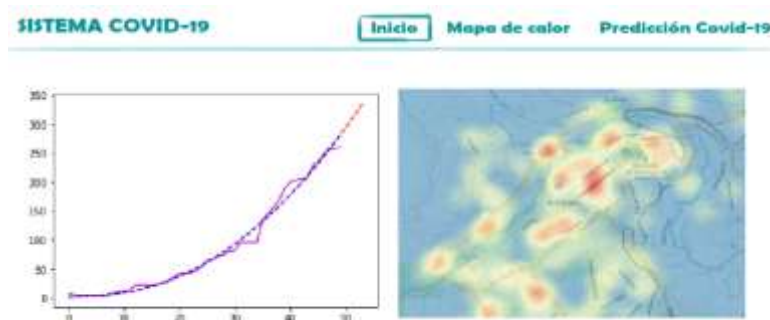


Gráfico 9 Diseño aplicación web
Fuente: El investigador.
Elaborado por: Kevin Estalin Llerena Ortiz.

En el grafico 9 podemos observar el diseño de la aplicación web integrando los dos módulos de predicción y de cerco epidemiológicos, siendo una interfaz sencilla, intuitiva y muy poderosa.

Tabla 8 Fase 3 Kanban
Fuente: El investigador.
Elaborado por: Kevin Estalin Llerena Ortiz.

Fecha: 12/12/2021		
Que hacer	Haciendo	Hecho
		Historia de usuario: • 1 • 2 • 3 • 4 • 5

En la tabla 8 podemos observar que el final de la fase 3 de Kanban, en el módulo “Hecho” ya se encuentran todas las historias de usuario y por ende el proyecto ha sido culminado con éxito.

ANÁLISIS DE RESTRICCIONES

El sistema necesitará datos en un archivo CSV (Valores separados por comas) en Excel caso contrario el algoritmo de predicción no funcionará.

El algoritmo para que obtenga mejor precisión se debe realizar predicciones máximo a 15 días caso contrario se obtendrá precisiones menores al 80%.

El cerco epidemiológico no tiene guías vectoriales solo son zonas de calor que ayudaran a delimitar las zonas de mayor riesgo.

ESTUDIO DE FACTIBILIDAD

Factibilidad Operativa

El sistema que se ha desarrollado en base a las necesidades y recomendaciones del departamento de tecnologías del honorable gobierno provincial de Tungurahua, poniendo en énfasis en que la interfaz sea fácil de manejar y con un diseño amigable, al mismo tiempo permitiendo al administrador generar reportes y realizar predicciones de forma inmediata de los contagios de COVID 19 en la provincia. Los beneficios que va a proporcionar el sistema a la institución son: La institución dispondrá de un algoritmo que predecirá contagios de la pandemia de covid-19 y también futuras pandemias con la cual se podrá tomar medidas rápidas y eficaces manteniendo la curva de contagios por debajo de la predicción. El sistema contará de un mapa de calor con el cual se podrá crear cercos epidemiológicos por parte de las autoridades competentes. El sistema podrá ser liberado a la ciudadanía o simplemente para manejos internos y uso de las autoridades de salud. El sistema generara reportes y así la toma de decisiones mejoraran sustancialmente. La aceptación del sistema cumple con las expectativas planteadas para ayudar en la toma de decisiones.

Factibilidad Técnica

La implementación de este proyecto con los demás módulos que conforman el sistema será en un servidor localhost. Al ser una aplicación web el sistema operativo del usuario es independiente, y en este caso se mantiene Windows 10 en los administradores del departamento de tecnologías.

Software

Para el desarrollo del proyecto se utiliza:

Herramienta de programación: Visual Studio Code utilizando como lenguaje de programación Python, Html, CSS con frameworks como Flask y Vuejs.

Complementos de desarrollo: QGIS

Hardware

La mayoría de las oficinas de que administran emergencias sanitarias del honorable gobierno provincial de Tungurahua cuentan con una computadora y están en óptimas condiciones para ejecutar sin dificultad el sistema.

Resultando que la factibilidad técnica no cuenta con ningún impedimento para la implementación del sistema.

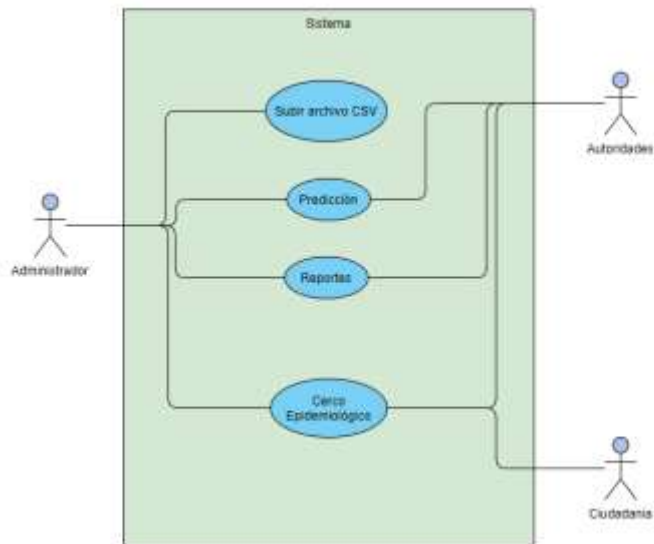
Factibilidad Económica

El proyecto desde un comienzo tuvo el apoyo incondicional del honorable gobierno provincial de Tungurahua proporcionando todo lo necesario en cuanto a equipos y servicios. Como herramienta de desarrollo se ha escogido Flask, Python, Vuejs y QGIS todos de software libre para evitar costo de licencias, de igual forma la base de datos esta implementada en un archivo CSV en Excel para evitar costes.

Obteniendo como resultado una aceptación total por parte de los administradores.

ANÁLISIS ORIENTADO A OBJETOS

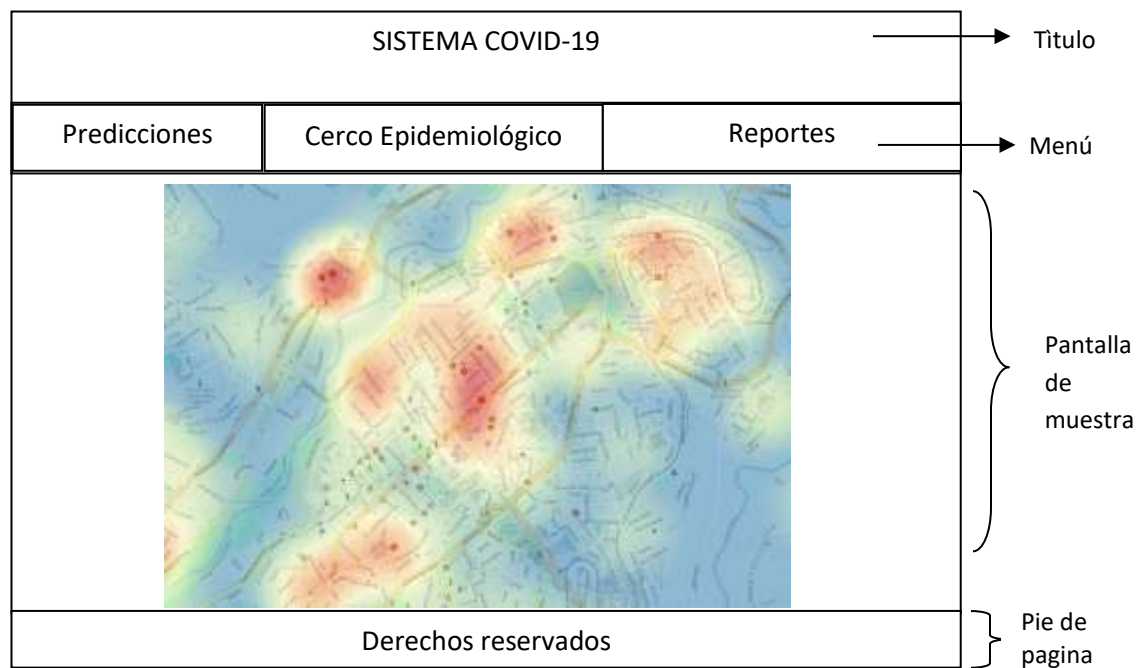
Diagramas de Casos de Uso



DISEÑO DE LA INTERFAZ DE USUARIO

Diseño de Entradas

Ventana 1



A través de esta ventana se podrá revisar el mapa de calor que indica los lugares que están en cerco epidemiológico.

CODIFICACIÓN

Código en con Python para predecir contagios Covid-19

Cargando datos:

```
#### CARGANDO DATOS ####
datos = pd.read_csv('Libro5.csv', sep=',')
datos = datos[['Dias', 'Casos']]
```

Preparando los datos:

```
#### PREPARANDO LOS DATOS ####
x = np.array(datos['Dias']).reshape(-1, 1)
y = np.array(datos['Casos']).reshape(-1, 1)
plt.plot(y, '-m')
polyFeat = PolynomialFeatures(degree=3)
x = polyFeat.fit_transform(x)
```

Entrenando los datos:

```
#### ENTRENANDO LOS DATOS ####
modelo = linear_model.LinearRegression()
modelo.fit(x,y)
precision = modelo.score(x,y)
y0 = modelo.predict(x)
```

Proceso de predicción:

```
#### PREDICCION ####
dias_prediccion = 5
x1 = np.array(list(range(1,50+dias_prediccion))).reshape(-1,1)
y1 = modelo.predict(polyFeat.fit_transform(x1))
plt.plot(y1, '--r')
plt.plot(y0, '--b')
```

PUESTA EN MARCHA

El sistema al ser una aplicación web se instaló todos sus componentes en el servidor local de la institución pudiendo ser puesta en marcha por las autoridades pertinentes, trabajando de forma local y a la vez dejando en potestad de la institución si el sistema fuere subido en un hosting para liberar al público en general, ya que al ser una aplicación web puede ser utilizado en cualquier navegador de internet.

RESULTADOS ESPERADOS

El Gobierno provincial de Tungurahua al servicio de la ciudadanía es exigente con los resultados de cada uno de sus proyectos es por ello que el sistema desarrollado para ayudar a controlar la curva de contagios de covid-19 en el provincia, cumple con todos los estándares de calidad y eficacia en sus propósitos; el sistema se espera que pueda ser utilizado para la generación de reportes y que dichos reportes ayuden a este problema de salud, que sea una herramienta de gran ayuda en la toma de decisiones que es lo más importante para frenar esta pandemia y ayudar a los sistemas de salud a no colapsar.

CAPITULO IV

CONCLUSIONES Y RECOMENDACIONES

CONCLUSIONES

- Luego de desarrollado el sistema de predicción de contagios y cerco epidemiológico del Covid-19 en la provincia de Tungurahua para el Honorable Gobierno Provincial de Tungurahua se puede concluir que será muy útil entre las autoridades pertinentes y las autoridades de salud para mejorar la toma de decisiones y así evitar el colapso de las casas de salud haciendo buen uso de las nuevas tecnologías en las cuales todo gobernante debería invertir como lo es el Machine learning con el cual podemos predecir futuras pandemias y estar mejor preparados.
- Se realizo un ETL (Extracción, Transformación, Carga) a la base de datos manual de casos COVID-19 en el cual se consiguió organizar los datos entregados por el ministerio de salud de tal manera que pudiesen ser utilizados por el algoritmo de predicción y la herramienta QGIS; dichos datos no pueden ser interpretados como se debería ya que no cuentan con un orden o protocolo para almacenar dichos datos y es por ello que el proceso ETL fue fundamental en el proceso del proyecto.
- Se aplico el algoritmo de regresión polinomial para predecir contagios COVID-19 en Python por ser uno de los lenguajes más poderosos y robustos para la aplicación del Machine learning. El correcto uso de estas tecnologías puede llegar a ser tan poderosas herramientas que facilitarían la toma de decisiones y por ende tener un mejor control ante estos acontecimientos de salud con carácter global.
- Se creo un mapa de calor de los casos COVID-19 con sus respectivas ubicaciones geográficas en QGIS para luego poder ser utilizado en la implementación de cercos epidemiológicos y así mantener a la ciudadanía

alerta y segura. Los mapas de calor no es algo nuevo, tampoco los cercos epidemiológicos, pero al tener un correcto uso de los datos y junto a los algoritmos de predicción se forma una herramienta robusta que ayuda a delimitar zonas de mayor riesgo mejorando la redistribución de los recursos para mitigar la curva de contagios.

RECOMENDACIONES

- A los entes públicos destinar más recursos en la parte tecnológica ya que sería de mayor ayuda en la toma de decisiones que realicen las autoridades competentes para salvaguardar la integridad de su población.
- A los entes de salud, almacenar y recolectar los datos de una manera organizado y siguiendo protocolos para evitar la fuga y redundancia de datos para que sean de fácil lectura y comprensión para futuras aplicaciones en la tecnología.
- Se recomienda utilizar las últimas tendencias de tecnologías y trabajar con ellas porque se puede obtener mejor soporte de la comunidad que utiliza dichas herramientas y tecnologías recomendando el lenguaje de programación Python para el machine learning y Django para la aplicación web, sin olvidar las buenas costumbres de programación.
- Es necesario que en la aplicación de machine learning se otorgue al programador las facilidades de acceso a los datos porque mientras más datos mejor se comportan los algoritmos de predicción.

LECTURA CITADA

- [1] «IEEE Xplore Full Text PDF». Accedido: jun. 17, 2021. [En línea]. Disponible en:
<https://ieeexplore.ieee.org/ielx7/6287639/8948470/09249010.pdf?tp=&arnumber=9249010&isnumber=8948470&ref=aHR0cHM6Ly9pZWVleHBsb3JlLmllZWUub3JnL2Fic3RyYWN0L2RvY3VtZW50LzkyNDkwMTA=>
- [2] «ScienceDirect Full Text PDF». Accedido: jul. 22, 2021. [En línea]. Disponible en:
<https://www.sciencedirect.com/science/article/pii/S1876034119302394/pdf?md5=d609e6c311170a4385ed57533c42a10d&pid=1-s2.0-S1876034119302394-main.pdf&isDTMRedir=Y>
- [3] «One Health Research Group, Facultad de Medicina, Universidad de las Américas, Quito, Ecuador. et al. - 2020 - Impacto de la COVID-19 en el Ecuador De los datos.pdf». Accedido: jul. 22, 2021. [En línea]. Disponible en:
<http://revecuatneurol.com/wp-content/uploads/2020/11/2631-2581-rneuro-26-02-00008.pdf>
- [4] «Reseña Histórica». Accedido: jul. 22, 2021. [En línea]. Disponible en:
<https://www.tungurahua.gob.ec/index.php/gobiernotungurahua/historia>
- [5] «Informática Básica: ¿Qué son las aplicaciones web?», *GCFGlobal.org*.
<https://edu.gcfglobal.org/es/informatica-basica/que-son-las-aplicaciones-web/1/> (accedido ene. 06, 2022).
- [6] R. PowerData, «¿Qué son los procesos ETL?» <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/qu-son-los-procesos-etl> (accedido ene. 06, 2022).
- [7] «What is ETL (Extract, Transform, Load)?», sep. 10, 2021.
<https://www.ibm.com/cloud/learn/etl> (accedido ene. 06, 2022).
- [8] «IBM Docs», mar. 08, 2021. <https://prod.ibmdocs-production-dal-6099123ce774e592a519d7c33db8265e-0000.us-south.containers.appdomain.cloud/docs/en/ztpf/1.1.0.15?topic=support-debug-information> (accedido ene. 06, 2022).
- [9] S. Yuill y H. Halpin, «Python». 2006.

- [10] «Google Books Link». Accedido: jul. 22, 2021. [En línea]. Disponible en: <https://books.google.com.ec/books?id=c8pV-TzyfBUC>
- [11] LUCA, «¿Qué es Python?», *¿Qué es Python?* <https://luca-d3.com/es/data-speaks/diccionario-tecnologico/python-lenguaje> (accedido jul. 22, 2021).
- [12] «Machine learning», *Wikipedia*. jul. 14, 2021. Accedido: jul. 22, 2021. [En línea]. Disponible en: https://en.wikipedia.org/w/index.php?title=Machine_learning&oldid=1033602835
- [13] «Análisis predictivo, qué es y para qué sirve», *BAOSS*, oct. 31, 2017. <https://www.baoss.es/analisis-predictivo-que-es/> (accedido jul. 22, 2021).
- [14] «ML-43: Algoritmo de regresión polinómica para aprendizaje automático + código fuente - programador clic». <https://programmerclick.com/article/1290232225/> (accedido jul. 22, 2021).
- [15] «QGIS», *Wikipedia, la enciclopedia libre*. jun. 22, 2021. Accedido: jul. 22, 2021. [En línea]. Disponible en: <https://es.wikipedia.org/w/index.php?title=QGIS&oldid=136507464>
- [16] «Heat map», *Wikipedia*. jun. 19, 2021. Accedido: jul. 22, 2021. [En línea]. Disponible en: https://en.wikipedia.org/w/index.php?title=Heat_map&oldid=1029382904
- [17] Diego, «Metodologías de Desarrollo Agile: Kanban Vs. Scrum Vs. XP», *Enjoy Life*, jul. 17, 2017. <https://enjoylife.com.ar/novedades/metodologias-desarrollo-agile-kanban-vs/> (accedido ago. 20, 2021).
- [18] «Scrum: pasado y futuro», *Netmind*, dic. 24, 2020. <https://netmind.net/es/scrum-el-pasado-y-el-futuro/> (accedido ene. 07, 2022).
- [19] «Metodología Kanban: en qué consiste y cómo utilizarla», *APD España*, jun. 08, 2021. <https://www.apd.es/metodologia-kanban/> (accedido ago. 06, 2021).
- [20] «Te damos 10 razones para usar la metodología Kanban en tu organización», *TecnoSoluciones.com*, jul. 14, 2020. <https://tecnosoluciones.com/10-razones-para-usar-la-metodologia-kanban-en-tu-organizacion/> (accedido ene. 07, 2022).

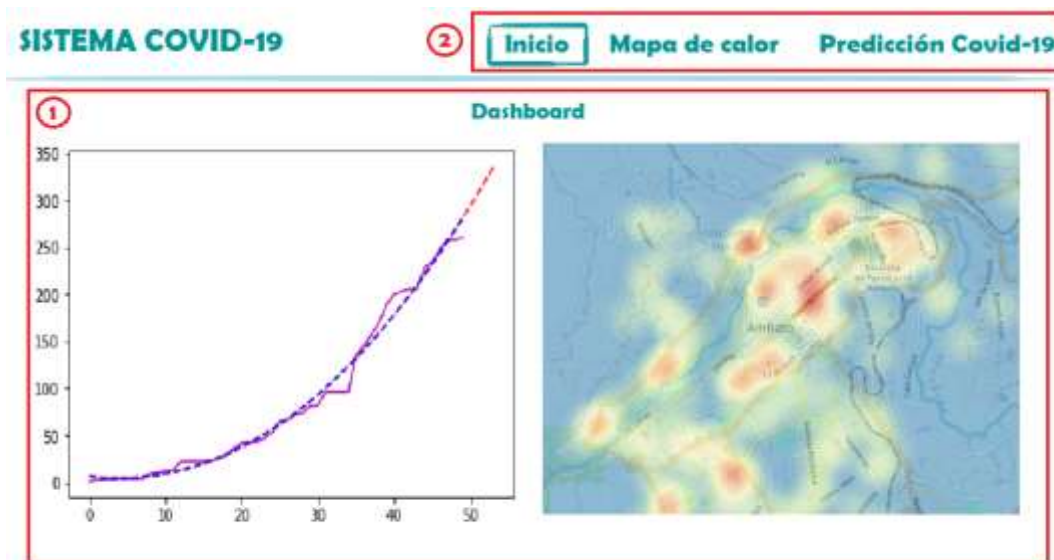
ANEXOS

MANUAL DE USUARIO SISTEMA COVID-19 TUNGURAHUA

Visión general de la aplicación web

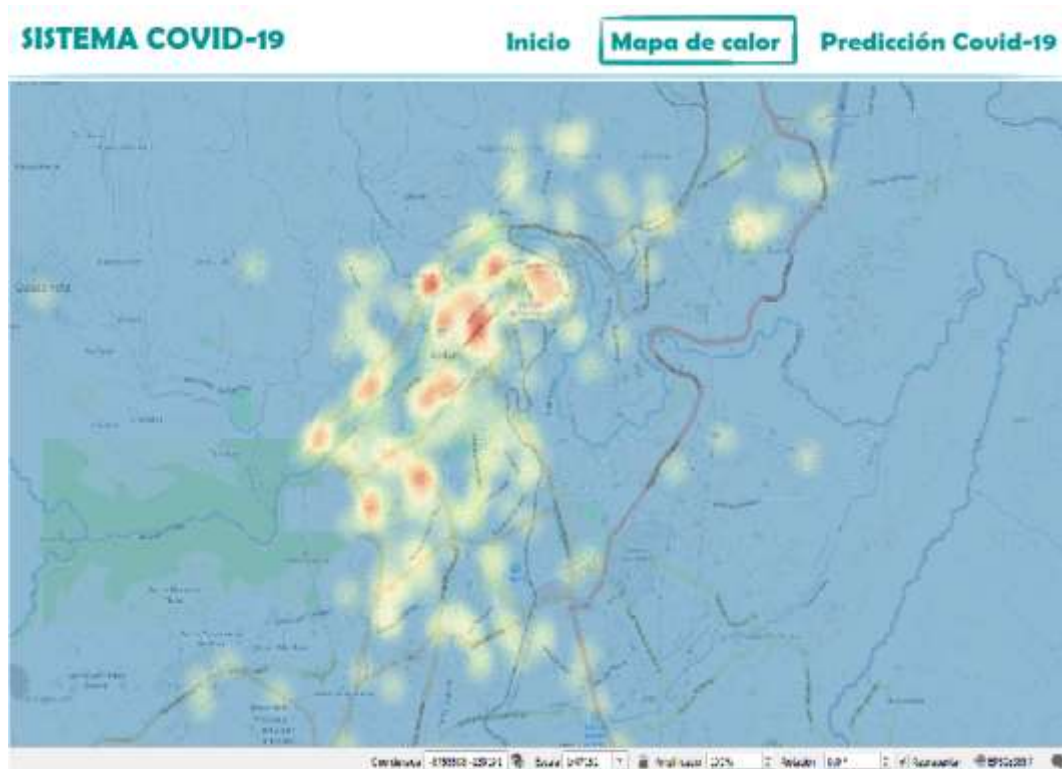
El panel principal consta de dos zonas principales.

1. **Dashboard:** Muestra la información generalizada en imágenes de los dos módulos (Predicción Covid-19 y Mapa de calor).
2. **Menú:** Muestra información de navegabilidad en la aplicación web como es el Inicio, Mapa de calor y Predicción Covid 19.



Paginas

Pagina Mapa de calor

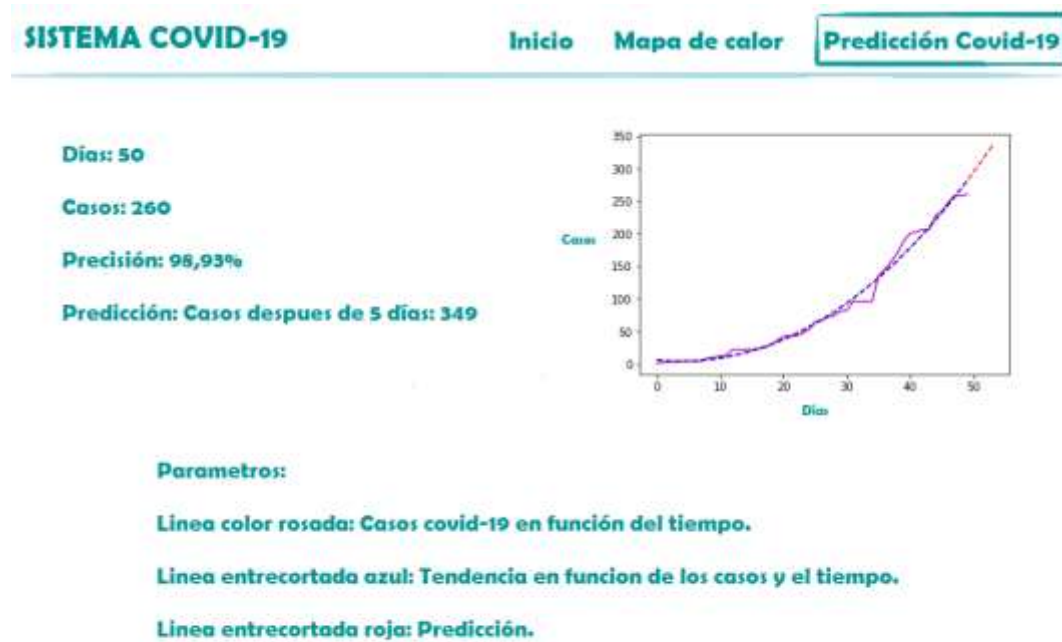


En esta pagina podremos observar el mapa de calor generado automáticamente con los datos que previamente han sido subidos al sistema, aquí se podrá visualizar y controlar con las herramientas que ofrecen un mapa geográfico digital.

Los parámetros a considerar son los siguientes:

- Color rojo: Zona de alto riesgo de contagio.
- Color Amarillo: Zona de riesgo moderado de contagio.
- Color Verde: Zona de bajo riesgo de contagio.
- Color Azul: Zona segura.

Pagina Predicción de contagios



En esta sección podemos observar la predicción a 5 días después de los primeros 50 días de haber confirmado el primer caso covid-19 en la provincia de Tungurahua, junto con sus datos y parámetros necesarios para realizar un análisis de parte de las autoridades competentes.

CODIFICACIÓN

Algoritmo de predicción de contagios covid-19 utilizando la regresión polinomial en Python.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.preprocessing import PolynomialFeatures
from sklearn import linear_model

#### CARGAR DATOS ####
data = pd.read_csv('Libro5.csv', sep=',')
data = data[['Dias', 'Casos']]
print('-'*30)
print('INICIO')
print('-'*30)
print(data.head())

#### PREPARAR DATOS ####
print('-'*30)
print('PREPARAR LOS DATOS')
print('-'*30)
x = np.array(data['Dias']).reshape(-1, 1)
y = np.array(data['Casos']).reshape(-1, 1)
plt.plot(y, '-m')
# plt.show()
polyFeat = PolynomialFeatures(degree=3)
x = polyFeat.fit_transform(x)
# print(x)
```

```
#### ENTRENAR DATOS ####
```

```
print('-'*30)
```

```
print('ENTRENAMIENTO DE DATOS')
```

```
print('-'*30)
```

```
model = linear_model.LinearRegression()
```

```
model.fit(x, y)
```

```
accuracy = model.score(x, y)
```

```
print(f'Precisión: {round(accuracy*100,2)} %')
```

```
y0 = model.predict(x)
```

```
#### PREDICCIÓN ####
```

```
days = 5
```

```
print('-'*30)
```

```
print('PREDICCIÓN')
```

```
print('-'*30)
```

```
print(f'Predicción – Casos despues de {days} dias: ', end="")
```

```
print(round(int(model.predict(polyFeat.fit_transform([[50+days]])))/1, 2), ")
```

```
#
```

```
print(round(int(model.predict(polyFeat.fit_transform([[234+days]])))/1000000,2),'Millio  
n')
```

```
x1 = np.array(list(range(1, 50+days))).reshape(-1, 1)
```

```
y1 = model.predict(polyFeat.fit_transform(x1))
```

```
plt.plot(y1, '--r')
```

```
plt.plot(y0, '--b')
```

```
plt.show()
```

ENTREVISTA

Modalidad: Virtual

Medio: Zoom

Tipo: Semi estructurada

Entrevistador: Kevin Llerena

Entrevistado: Ing. Galo Robayo

Pregunta:	Respuesta:	Interpretación:
#1: ¿Cuáles cree usted que son las necesidades que tiene la prefectura y las autoridades de emergencia para enfrentar la pandemia del covid-19?	Yo creo que las autoridades del COE Provincial ante tal situación que atraviesa el país necesitan la ayuda de la tecnología para sobrellevar esta emergencia sanitaria y así mejorar la toma de decisiones con ayuda de un sistema que prediga los contagios y realice cercos epidemiológicos.	El jefe del departamento de tecnologías siente que las autoridades de emergencia necesitan manejar los datos con nuevas tecnologías que ayuden en la toma de decisiones.
#2: ¿Como y en que formato recogen los datos de contagios de covid-19 en la provincia o el ministerio de salud les proporciona estos datos para que las autoridades puedan tomar decisiones?	Los datos son confidenciales y nos proporciona el ministerio de salud, están guardados en un Excel y solo contienen las direcciones de los contagiados.	Los datos los manejan de forma manual y utilizan una hoja de cálculo en la cual no hay latitud ni longitud que se necesitarían para realizar los cercos epidemiológicos con mapas digitales.

#3: ¿Existe alguna manera en la que utilicen el machine learning para predecir futuros contagios de covid-19 en la provincia?	En el departamento de tecnologías del consejo provincial no existe algo similar y simplemente no se puede generar reportes para las autoridades todo se analiza manualmente.	Los datos los analizan manualmente por ende no existe un patrón a seguir por parte de las autoridades lo que afectaría de forma negativa en la toma de decisiones.
#4: ¿Qué tecnologías me recomendaría utilizar para el desarrollo del sistema de predicción de contagios y cercos epidemiológicos?	La verdad te recomendaría utilizar Matlab para realizar las predicciones con algún algoritmo de machine learning y para los mapas de cercos epidemiológicos utilizar ArcGIS y eso mostrar en una interfaz web.	Se deberá analizar con que tecnologías trabajar para el desarrollo del sistema priorizando software libre.
#5: ¿La información brindada por el sistema se compartirá con la población?	En cierto punto sí, pero eso depende de las políticas de las autoridades y la confidencialidad para manejos internos.	La necesidad de mostrar tal información en la web era una inquietud para analizar en que se desarrollara y como se mostraran las predicciones y cercos epidemiológicos.
#6: ¿Se necesitará la generación de reportes?	Claro que sí, la generación de reportes es muy necesario para que las autoridades puedan tomar decisiones basadas	La generación de reportes será algo indispensable en el sistema a desarrollar.

	en las predicciones otorgadas por el sistema.	
--	--	--



Ambato abril 21, 2022



CERTIFICADO

A quien interese

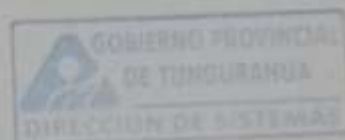
Yo, Galo Xavier Robayo Laz, con número de cédula 0503161770, director de la Dirección de Sistemas del Gobierno Provincial de Tungurahua, notifico la **ACEPTACIÓN** del proyecto **"DESARROLLO DE UN SISTEMA DE PREDICCIÓN DE CONTAGIOS Y CERCO EPIDEMIOLÓGICO DEL COVID-19 EN LA PROVINCIA DE TUNGURAHUA PARA EL HONORABLE GOBIERNO PROVINCIAL DE TUNGURAHUA"** que se llevó a cabo por el señor Llerena Ortiz Kevin Estalin a partir del 22 de septiembre de 2020 al 16 de octubre de 2021.

Como se indicó en la presentación del proyecto, el señor Kevin Llerena estará a cargo del desarrollo e implementación del proyecto y esta Dirección será responsable del buen uso del sistema.

La culminación de este proyecto se llevó a cabo bajo las condiciones y características estipuladas en la convocatoria.

Atentamente,

Ing. M.Sc. Galo Robayo Laz
Director Dirección de Sistemas



GRUdi