

Sistema preventivo de contagio de COVID-19 en ambientes empresariales con uso de técnicas de clasificación de Minería de datos.

Trabajo Terminal No. 2021-A068

*Alumnos: *Hernandez Cruz Cristian, Torres Miranda Christian*

Directores: Araujo Diaz David, Méndez Segundo Laura

chernandezc1705@alumno.ipn.mx

Resumen: Este sistema pretende, mediante la obtención de datos particulares de un usuario, informar de la presencia del virus COVID-19 en las instalaciones de su empresa, en estrecha relación con datos como los síntomas recientes, padecimientos crónicos y otros datos que ayudaran al sistema a dar una respuesta con ayuda de algunas estrategias de minería de datos.

Palabras clave: Sistema web, minería de datos, bases de datos.

1. Introducción

En la actualidad la pandemia se ha visto mitigada enormemente, con la mayoría de los estados del país en semáforo verde y con la Ciudad de México a punto de unirse a estos. A pesar de esto, la contingencia aún no ha acabado y es importante mantenerse informado y seguir acatando las medidas sanitarias para volver a una “normalidad”. Afortunadamente las circunstancias permiten que cada vez más se regrese a trabajar de forma presencial tanto en el ámbito escolar como el comercial. Pero esta situación le da la sensación a una gran parte de la población de que ya se puede salir a las calles a retomar una vida tal como era antes del inicio de la pandemia, lo cual es incorrecto y debe de corregirse aportando información, ya que existen aún sitios en donde es potencialmente probable ser contagiado, Este sistema ataca directamente esta problemática, ya que la intención es dar a las empresas información personalizada acerca de su situación respecto a la pandemia, a comparación de la información que se puede consultar en medios públicos como televisión, notas periodísticas o páginas de estadísticas generales. Además, cuenta con la característica de enfocarse principalmente en ambientes empresariales, en donde las zonas de trabajo se dividen en departamentos y es importante tener un seguimiento del personal y los usuarios que laboran ahí. El análisis se realizará mediante un conjunto de datos obtenidos de un registro diario de los usuarios en el que se contestaran un conjunto de preguntas claves para conocer su estado actual, a estos datos se les aplicaran un conjunto de procesos de minería de datos [1] para obtener un resultado.

Existen algunos sistemas desarrollados en esta pandemia que aportan una solución e información a situaciones similares a las cuales pretende ayudar este sistema, algunos de ellos son:

1. #juntosvenceremoscovid19[2]
2. ZYGHT|COVID-19[3]
3. COVID-19MX[4]
4. Business Process Studio for Jira[5].

Las propuestas anteriores se describen a continuación:

SOFTWARE	CARACTERÍSTICAS	PRECIO EN EL MERCADO
#juntosvenceremos covid19	Aplicación que permite conocer, de forma individual, su riesgo de contraer el virus, así como un mapeo de las zonas de mayor riesgo de contagio.	Software gratuito perteneciente a la UNAM.
ZYGHT COVID-19	Software para el seguimiento del cumplimiento de las medidas preventivas en empresas e industrias consideradas indispensables.	Basado en cotización
COVID-19MX	Aplicación móvil que permite realizar un autodiagnóstico, concede acceso directo al teléfono de atención epidemiológica sanitaria e identifica los Centros de Atención más cercanos a tu ubicación.	Aplicación móvil gratuita disponible en la Google Playstore y en la Appstore.
Business Process Studio for Jira	Modulo para Atlassian Jira que permite la generación rápida y visual de los procesos necesarios para el seguimiento de los contagios sin tener que escribir código.	Basado en cotización
Solución propuesta	Software de seguimiento de contagios en empresas dependiendo de su situación interna y externa a la empresa, generación de reportes diarios, posibilidad de reportar un posible contagio.	Software perteneciente al IPN

Tabla 1. Resumen de productos similares

2. Objetivo

Desarrollar un sistema orientado a empresas con una gran cantidad de empleados que permita, tomando como parámetros datos proporcionados por la plantilla laboral, determinar zonas de alto contagio en las instalaciones, dando la posibilidad de actuar a tiempo y prevenir un contagio masivo.

3. Justificación

El problema del regreso a oficinas de forma presencial es lo delicado que resulta que un miembro de la empresa sea portador del virus, lo cual resultaría en un nuevo cierre temporal de actividades presenciales. Este sistema tiene como principal función hacer uso de tecnologías de minería de datos para prevenir un posible contagio en alguna de las áreas de trabajo y así evitar una situación más grande y grave.

Se han realizado aplicaciones o sistemas con una propuesta similar, notificar de manera personal al usuario de su situación respecto al virus en base a sus síntomas y circunstancias como su entorno. Lo que hace diferente a nuestro sistema es la posibilidad de convertir estos resultados obtenidos de cada usuario en una herramienta útil

para empresas de un tamaño considerable, en donde una gran cantidad de personas laboran diariamente, por lo que se vuelve complicado tener una certeza del estado de salud de cada uno de ellos.

Además, la minería de datos nos permitirá entregar resultados que nos ayuden a determinar zonas de posible alto contagio antes de la tragedia, y así evitar un cierre precipitado de las instalaciones por culpa de un contagio.

Las dos técnicas de minería de datos que se utilizarán de manera más probable en el sistema son el clustering y los árboles de decisión ID3, aunque esto se analizará de manera más exhaustiva en el proceso.

La primera nos permitiría conocer por grupos a las personas que comparten características que nos interesen conocer, como el grupo de empleados que tienen un padecimiento crónico en concreto, las personas que sean clasificadas en este grupo serían consideradas con una mayor probabilidad de portar el virus y contagiarlo en las instalaciones de la empresa.

La segunda técnica sería de utilidad para determinar zonas de alto contagio dentro de la propia empresa, ya que dependiendo de la información que se nos proporcione por medio del sistema, se podrá determinar la probabilidad de que sea contagiado fuera de la empresa, y si se detectan múltiples individuos con la misma situación, se puede determinar ese departamento o área de la empresa como una zona de alto contagio, en donde se indicará dentro del sistema que se deberán tomar medidas más estrictas para evitar un cierre de la misma.

Las variables que se utilizarán para el análisis son datos proporcionados por los empleados como preguntas comunes acerca de síntomas recientes, contacto cercano reciente con un posible portador del virus, entre otras preguntas que nos ayudaran a determinar el riesgo de que el sujeto sea contagiado.

Al final se contará con un resumen de cada uno de los departamentos de la empresa, con su respectiva información acerca de la posibilidad de que un contagio ocurra ahí.

Se utilizarán distintas herramientas y conocimientos adquiridos a lo largo de la carrera, como la antes mencionada minería de datos, conocimientos en sistemas web y bases de datos, entre otras áreas de conocimiento que nos permitirá entregar un resultado satisfactorio al final de la evaluación, con un sistema web amigable con el usuario y útil para el público objetivo.

4. Productos o resultados esperados

La arquitectura del sistema y el proceso de entrega del reporte final se muestran en el siguiente diagrama de bloques:

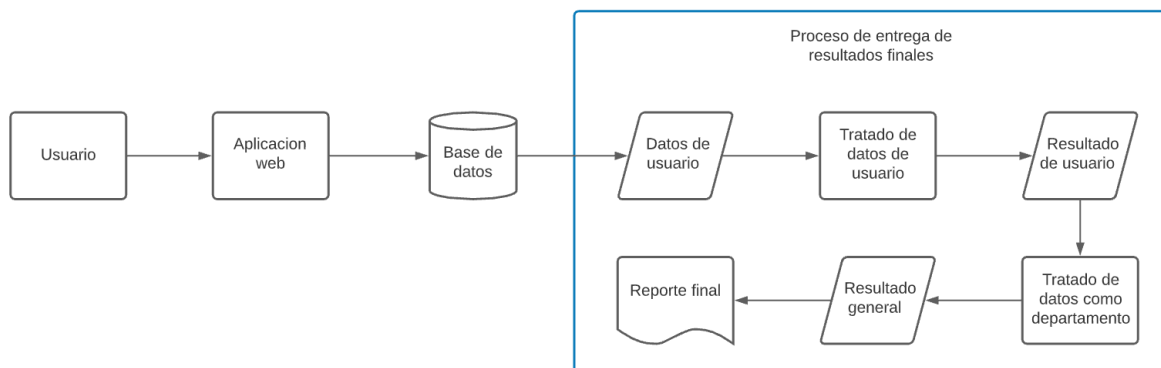


Figura 1. Arquitectura del sistema

Las tecnologías que servirán de apoyo para la creación de este sistema web son las siguientes:

- HTML5
- CSS
- Django 3.2
- MySQL 8.0
- Knime

El sistema tiene contemplados dos tipos de usuarios:

- **Trabajador:** Puede realizar el registro en la aplicación web de manera diaria y también puede reportar una falta al administrador. En caso de emergencia puede realizar un reporte de contagio.
- **Administrador:** Puede configurar los distintos departamentos de la empresa para su posterior análisis, recibe los resultados de cada departamento resumido en el reporte final.

El proceso inicia desde la configuración de una nueva empresa, ya que la aplicación otorga la capacidad y flexibilidad de configuración de la empresa dependiendo de su tamaño, número de trabajadores registrados y número de departamentos.

Con la empresa configurada y añadida a la aplicación, se inicia el proceso de predicción de resultados:

1. El usuario se registra a la aplicación, y con una cuenta creada solamente registra los datos necesarios diariamente.
2. El servidor recibe los datos y los envía al módulo de tratado de datos del usuario, en donde se aplican las técnicas de minería de datos pertinentes.
3. Los datos resultantes son enviados al módulo de tratado de datos por departamento, repitiendo el proceso anterior.
4. Se genera un reporte con los resultados finales.

Los productos que se entregaran al final de desarrollo son:

1. Aplicación web que permita acceder como “trabajador” y “administrador”.
2. Código fuente del proyecto.
3. Reporte técnico.
4. Exposición de resultados ante el jurado.

5. Metodología

La metodología elegida para el desarrollo de este sistema es el modelo V[6]. Esto es debido al tamaño del proyecto y del tamaño del equipo de desarrollo, y al ser este un proyecto desarrollado por dos personas es la metodología adecuada.

Este modelo se divide en las siguientes fases:

- Fase de especificaciones (se establece el objetivo del sistema y la problemática a resolver).
- Fase funcional (se determina de manera teórica las funciones y capacidades del sistema).
- Fase de diseño (se analiza la arquitectura del sistema y de los distintos módulos que lo conformaran).
- Codificación (se realiza la programación del sistema).
- Pruebas de unidad (se realizan las pruebas que satisfagan a la fase de diseño).
- Pruebas de integración (se realizan las pruebas que satisfagan a la fase funcional).
- Integración del sistema (se realizan las pruebas que satisfagan a la fase de especificaciones).
- Validación (se confirma que el sistema cumple con la calidad y objetivos establecidos).

6. Cronograma

Nombre del alumno(a): Hernandez Cruz Cristian

Título del TT: Sistema predictivo de contagio de COVID-19 en ambientes empresariales con uso de Minería de Datos.

[illegible]

Presentación de resultados											
Evaluación de TT II.											

7. Referencias:

[1] Ana Gonzales Marcos. (2006). Desarrollo de técnicas de minería de Datos en procesos industriales. 2021, de Universidad de la Rioja Sitio web: http://repositorio.usil.edu.pe/bitstream/USIL/2763/1/2017_Granda_Aplicacion-de-mineria-datos.pdf

[2] Laboratorio de Instrumentación Espacial (LINX). (2020). #juntosvenceremoscovid19. 2021, de UNAM Sitio web: <https://covid19.sol21ct.com/support>.

[3] Prevencionar. (2020). ZYGHT|COVID-19 – software para el seguimiento de las medidas preventivas. 2021, de Prevencionar Sitio web: <https://prevencionar.com/2020/04/27/zyghtcovid-19-software-para-el-seguimiento-de-las-medidas-preventivas/>

[4] Sistel. (2020). SISTEL IMPLANTA EL SEGUIMIENTO DE CASOS COVID-19 EN LA EMPRESA CON BUSINESS PROCESS STUDIO FOR JIRA. 2021, de Sistel Sitio web: <https://www.sistel.es/sistel-implanta-seguimiento-casos-covid-19-empresa-business-process-studio-jira>

[5] Secretaria de Salud MX. (2020). COVID-19MX. 2021, de Secretaria de Salud MX Sitio web: https://play.google.com/store/apps/details?id=mx.gob.www&hl=es_MX&gl=US

[6] Roger S. Pressman. (2010). Desarrollo ágil. En Ingeniería de software: Un enfoque practico (67-76). México: McGraw Hill.

8. Alumnos y directores

Hernandez Cruz Cristian.- Alumno de la carrera de Ingeniería en Sistemas Computacionales en ESCOM, Boleta: 2018630231, Tel. 7757525617, Email: chernandezc1705@alumno.ipn.mx.

Torres Miranda Christian – Alumno de la carrera de Ingeniería en Sistemas Computacionales en ESCOM, Boleta: 2014090716, Tel. 5534386592, Email: ctorresm1301@alumno.ipn.mx.


Araujo Díaz David.- Profesor de la Escuela Superior de Cómputo (Depto. de Posgrado), Ing. en Comunicaciones y Electrónica (ESIME-Zacatenco), M. en C. en Ingeniería Eléctrica en la Opción de Computación (CINVESTAV-IPN). Áreas de Interés: Realidad Virtual, Diseño y Simulación de Circuitos Electrónicos, Modelación Matemática, Cómputo Paralelo, Robótica, Procesamiento de Imágenes, Reconocimiento de Patrones, Computabilidad, Complejidad Algorítmica y Seguridad Informática. Tel: 57-29-60-00 Ext. 52038. Email: daraujo@ipn.mx

Méndez Segundo Laura.- M. en C. con especialidad en computación CINVESTAV 1998, Licenciatura en Informática, Universidad Veracruzana 1991. Certificado de SCRUM Master en el 2017, Áreas de Especialidad: Bases de Datos, Ingeniería de Software, UML. Área de Interés cómputo educativo y procesamiento de imágenes. Tel: 57-29-60-00 Ext. 52032, Email: lmendez@ipn.mx

CARÁCTER: Confidencial
FUNDAMENTO LEGAL: Artículo 11
Fracc. V y Artículos 108, 113 y 117 de la
Ley Federal de Transparencia y Acceso a
la Información Pública.
PARTES CONFIDENCIALES: Número
de boleta y teléfono.

Acuses de recibo


M. en C. David Araujo Diaz:



Cristian Hernandez Cruz
Para: David Araujo Diaz

Protocolo.docx
86 KB

Buen día profesor le envié el protocolo del TT a espera de su acuse de recibo. Muchas gracias. Saludos.




David Araujo Diaz
Para: Cristian Hernandez Cruz

Recibido, ¡muchas gracias!

...

Responder Reenviar


M. en C. Laura Méndez Segundo:



Cristian Hernandez Cruz
Para: Laura Mendez Segundo

Protocolo.docx
87 KB

Buen día profesora. Le agradezco las observaciones, había varios puntos que tenían que ser corregidos. Le mando el protocolo con las correcciones incluidas. Espero su acuse de recibo y aprobación final. Saludos.



Laura Mendez Segundo
Para: Cristian Hernandez Cruz

Cristian:

RECIBÍ PROTOCOLO Y ESTOY DE ACUERDO.

Saludos cordiales

Laura Méndez Segundo


...

Muchas gracias. Gracias. Gracias por confirmar.

¿Las sugerencias anteriores son útiles? ☒ Sí ☐ No

Responder Reenviar

Christian Torres Miranda:



Christian Torres Miranda
Para: Cristian Hernandez Cruz

Recibí el protocolo con tema "Sistema preventivo de contagio de COVID 19 en ambientes empresariales con uso de técnicas de clasificación de Minería de Datos" y estoy de acuerdo con el registro del mismo.

Responder Reenviar