Inspecciones.

Informe.

Escuela de Informática y Telecomunicaciones

Big Data

# **Introducción.**

El presente informe tratará sobre la problemática de las viviendas con conversiones ilegales[[1]](#footnote-1), y así intentar generar un sistema que permitiera predecir que quejas deberían ser atendidas con la mayor urgencia.

# **Descripción del Problema.**

En la ciudad de Nueva York no hay suficientes inspectores para que inspeccionen las viviendas con “conversiones ilegales”, además, los inspectores investigan las quejas que les parecen de peor agüero[[2]](#footnote-2) y no en base a un estudio estadístico.

No parecía haber ningún sistema para distinguir los casos meramente molestos, de los que estaban a punto de estallar.

# **Descripción de la Solución Propuesta.**

Solución caso.

Se propuso usar muchos datos, agregar conjuntos de datos de 19 organismos distintos que indican cosas como por ejemplo, si el propietario no paga impuestos inmobiliarios, si tiene ejecuciones hipotecarias o si el cese de pago había causado un corte en el servicio eléctrico. También se incorporó información sobre la infraestructura del edificio, amén de ambulancias, tasa de delitos, quejas por roedores, etc. Luego se comparó la información con otros 5 años sobre incendios clasificados por gravedad y buscaron correlaciones, con esto se esperaba formar un sistema que logrará predecir quejas que deberían ser atendidas con urgencia.

Inventaron un sistema de identificación por edificio, se salió a terreno para investigar a los inspectores, se les hicieron preguntas y tomaron nota. Se percataron que los inspectores evaluaban según si el edificio se veía renovado o no, entonces se introdujo el dato “renovado reciente”.

*“Construir un equipo con los mejores científicos de datos que pudieran encontrar y explotar los montones de información virgen de la ciudad para aumentar la eficiencia en todos los terrenos posibles.”* (Instituto Profesional Duoc UC).

# **Elección y fundamentación de la Arquitectura que mejor se ajusta para solucionar el problema.**

Se propone el uso de la arquitectura kappa, es la que mejor se adapta a este caso debido a que se estarán cargando datos por lotes lo que nos garantiza mayor flexibilidad y esto nos permite un mayor soporte para el manejo y manipulación de la información asincrónica ya que no se procesan datos activamente.

# **Estrategia para garantizar una alta disponibilidad de los servicios.**

* Inversión en infraestructura como switches, servidores, circuitos de internet y storage
* Planificación exhaustiva en la implementación de la conectividad y redes
* Integración correcta entre los distintos componentes del sistema
* Entrega de hardware correspondiente a las distintas estaciones de trabajo (los pc de los operadores contarán con requisitos mínimos)

# **Políticas de respaldo de datos.**

Cada día se realizará un respaldo de emergencia de los datos en un servidor diferente al que está en operación.

Recuperación de desastres, que significa superar las contingencias que se puedan producir, independientemente de su origen y está sustentada por un plan de recuperación, cuyo objetivo es proporcionar los medios alternos para realizar las funciones normales, cuando los medios habituales no están disponibles a una contingencia.

# **Estrategia de seguridad.**

Habrá varias características para la seguridad:

* Inicio de sesión: El usuario autorizado, tendrá acceso a los datos con un correo electrónico y contraseña.
* Control de acceso basado en el rol: El acceso a los trabajos, tablas de datos y recursos.
* Paso a través de credenciales.
* Listas de acceso IP: Exija la ubicación de red de los usuarios del área de trabajo

# **Prueba de concepto en la cual se generé la carga de datos desde una fuente origen a Hadoop y Spark y se aprovechen ejecutando un caso de uso.**

# **Conclusiones.**

Para concluir creemos firmemente que nuestro sistema responde de manera efectiva a todas las necesidades planteadas en el caso, nuestro sistema es capaz de predecir de manera eficaz cada uno de los posibles escenarios, específicamente los de mayor catástrofe para así lograr una intervención lo más rápida posible.

# **Bibliografía.**

Instituto Profesional Duoc UC. (s.f.). Situación Evaluativa al estudiante. *ET121\_3\_BDY7101*. Santiago, Chile. Recuperado el 09 de Junio de 2021, de https://drive.google.com/drive/folders/1RygwbLnPl54TswqZmnUh5qd5YawDZTMy

Microsoft. (11 de Junio de 2021). Recuperado el 26 de Junio de 2021, de Microsoft: https://docs.microsoft.com/es-es/azure/databricks/security/security-overview-azure

Real Academia Española. (s.f.). *Real Academia Española*, 23.4. Recuperado el 9 de Junio de 2021, de RAE: https://dle.rae.es/ag%C3%BCero

**Contenido.**

[**Introducción.** 1](#_Toc74147446)

[**Descripción del Problema.** 1](#_Toc74147447)

[**Descripción de la Solución Propuesta.** 1](#_Toc74147448)

[**Elección y fundamentación de la Arquitectura que mejor se ajusta para solucionar el problema.** 2](#_Toc74147449)

[**Estrategia para garantizar una alta disponibilidad de los servicios.** 2](#_Toc74147450)

[**Políticas de respaldo de datos.** 3](#_Toc74147451)

[**Estrategia de seguridad.** 3](#_Toc74147452)

[**Prueba de concepto en la cual se generé la carga de datos desde una fuente origen a Hadoop y Spark y se aprovechen ejecutando un caso de uso.** 3](#_Toc74147453)

[**Conclusiones.** 3](#_Toc74147454)

[**Bibliografía.** 3](#_Toc74147455)

1. Conversiones ilegales: Práctica de subdividir un alojamiento en muchas unidades más pequeñas para acabar acomodando hasta 10 veces más personas que lo proyectado. [↑](#footnote-ref-1)
2. Agüero: *“Procedimiento o práctica de adivinación utilizado en la Antigüedad y en diversas épocas por pueblos supersticiosos, y basado principalmente en la interpretación de señales como el canto o el vuelo de las aves, fenómenos meteorológicos, etc.”* (Real Academia Española, s.f.). [↑](#footnote-ref-2)