



Facultad de Ingeniería

Escuela de Ingeniería en Bioinformática

Covid-19 en Chile y modelo epidemiológico SEIR

Taller de desarrollo de software 2020-2

Alumno:
Claudio Quevedo

Profesor:
Alejandro Valdés

Índice

1. Introducción	3
2. Contexto general	4
2.1. Cliente	4
2.2. Problema o necesidad	4
2.3. Solución	4
3. Metodología	6
4. Desarrollo	8
4.1. Requerimientos	8
4.2. Diseño	8
4.2.1. Base de datos	9
4.2.2. Tecnologías	9
4.3. Implementación	9
4.3.1. Iteración 1, descarga y procesamiento de información.	9
4.3.2. Iteración 2, programación de scripts para calcular el modelo SEIR y guardar sus resultados.	10
5. Conclusiones	11

1. Introducción

En este informe se presenta el desarrollo de un software en base al modelo epidemiológico matemático SEIR (susceptible, expuesto, infectado, recuperado) en las comunas de Chile con la diferenciación entre aplicar distanciamiento social y no aplicarlo. Esta software utilizara las bases de datos otorgadas por el Ministerio de Ciencia y Tecnología junto a los datos entregados por el Ministerio de Salud de Chile, además de parámetros otorgados por el Colegio Médico de Chile. Posterior a los cálculos realizados por el modelo, se aplicarán algoritmos de minería de datos e inteligencia artificial para la clasificación de las comunas en grupos de necesidad de recursos médicos (que pueden ser de mayor o menor necesidad). El propósito del software es la ayuda a la toma de decisiones para la distribución de implementos y de personal médico en las distintas localidades del país en base a las necesidades de éstas.

A su vez se hará uso de metodologías ágiles para su desarrollo. El desarrollo ágil de software se refiere a metodologías de desarrollo de software centradas en la idea de desarrollo iterativo, donde los requisitos y las soluciones evolucionan a través de la colaboración entre equipos multifuncionales auto-organizados. El valor máximo en el desarrollo ágil es que permite a los equipos entregar valor más rápido, con mayor calidad y predictibilidad, y con mayor aptitud para responder al cambio. XP (eXtreme Programming) es una metodología ágil, fundada por Kent Beck, centrada en potenciar las relaciones interpersonales como clave para el éxito en desarrollo de software, promoviendo el trabajo en equipo, preocupándose por el aprendizaje de los desarrolladores, y propiciando un buen clima de trabajo.

2. Contexto general

2.1. Cliente

Chile también se a visto afectado por la pandemia mundial de covid-19, por lo cual el cliente objetivo es aquel que debe tomar las desiciones sobre como y donde se enviarán los recursos médicos disponibles en el país.

A su vez como será un servicio web, cumplirá un rol de servicio social para aquellas personas que quieran saber el estado de las predicciones en su propia comuna.

2.2. Problema o necesidad

El pasado 3 de Marzo de 2020 se confirmó el primer caso de Covid-19 en Chile, hoy 28 de Noviembre ya existen 548.941 casos confirmados¹. Actualmente en Chile no hay ningún software que ayude a la toma de decisiones respecto a la distribución de recursos médicos basados en un análisis predictivo, en tiempo real (entregándole los datos al modelo) y personalizado para cada región o comuna del país.

El estudio de las enfermedades infecciosas a menudo se basa en modelos epidemiológicos matemáticos que intentan simular la dinámica de la enfermedad y estimar los parámetros relacionados con ella, como la tasa de reproducibilidad, velocidad de transmisión, fuerza de acción del gobierno, intensidad de reacción individual, entre otras.

Un tipo de simulación es el modelo SEIR basado en el supuesto de que la población se puede clasificar en cuatro grupos compartimentados independientes (individuos susceptibles, individuos expuestos, individuos infectados e individuos recuperados). Este modelo estudia de qué manera los individuos pueden progresar de un grupo compartimentado al siguiente.

2.3. Solución

Desarrollo de un software (servicio web) que utiliza el modelo SEIR para estudiar de qué manera los individuos pueden progresar de un grupo compartimentado al siguiente, esto ayudara a la toma de decisiones, tanto para estimar la necesidad de recursos médicos de

¹<https://www.minsal.cl/nuevo-coronavirus-2019-ncov/casos-confirmados-en-chile-covid-19/>

regiones y comunas, como la comparación entre efectuar y no efectuar un distanciamiento social.

El cual como input tiene el listado de cada comuna y su respectiva cantidad de habitantes, cantidad de infectados, cantidad de recuperados y su tasa de migración (o movimiento de personas entre comunas).

Luego la realización de los cálculos con el modelo SEIR y algoritmos de minería de datos e inteligencia artificial. Y finalmente el output será entregado en el servicio web, en el cual como se puede ver en la Figura 1, se extraerá los máximos de cada curva de infectados con y sin cuarentena de cada comuna.

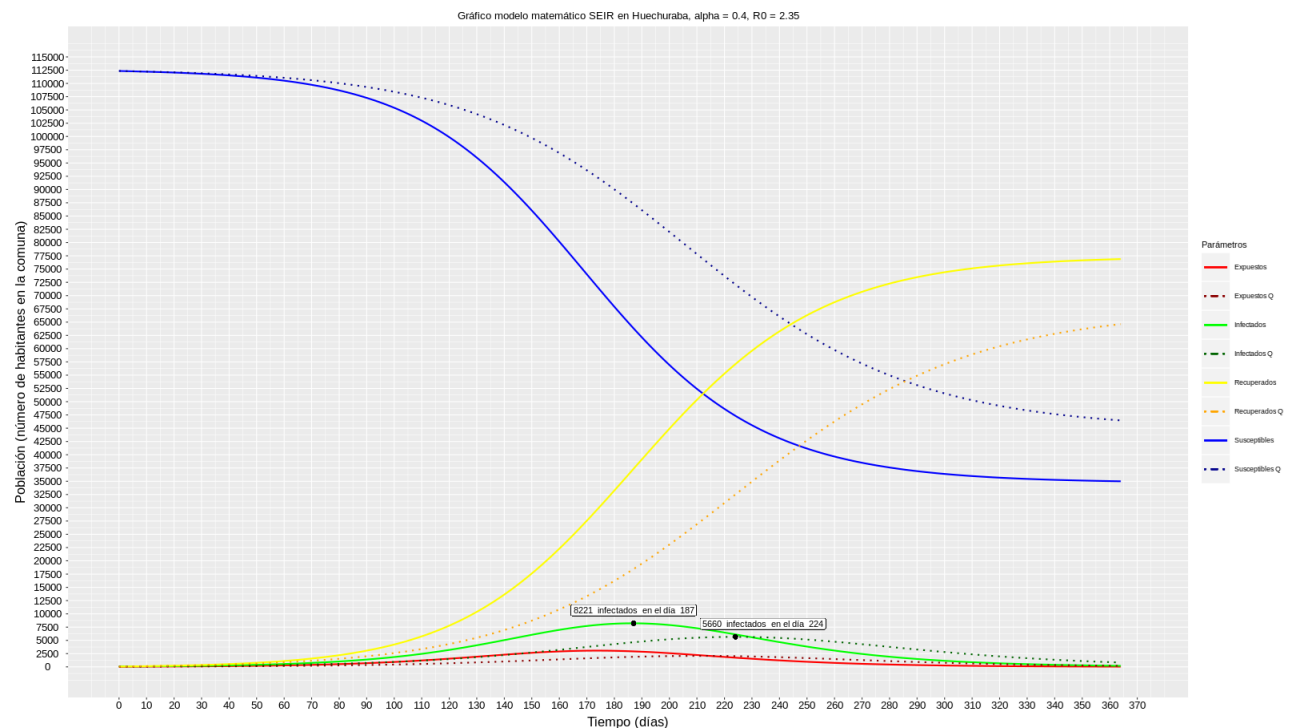


Figura 1: Ejemplo de gráfico del modelo epidemiológico matemático SEIR en la comuna de Huechuraba. En el eje “y” se encuentra la población, en el eje “x” los días (un año). En la cual se observan las curvas lineales que representan sin distanciamiento social y las curvas punteadas con distanciamiento social (en la leyenda inicia con “Q” haciendo referencia a la palabra en inglés “quarantine”). Uno de los análisis importantes del gráfico es que al hacer distanciamiento social ralentizaría (en 37 días) la velocidad de contagio, lo que ayudaría a no sofocar el sistema de salud. Otro análisis importante es el máximo de infectados, que se traduce a menos 2561 personas infectadas de las que se tendrán sin distanciamiento social.

3. Metodología

Se utilizará la metodología ágil XP (programación extrema), el cual es uno de los marcos de desarrollo de software más importantes². Así lograr mejorar la calidad del software y responder a los requisitos del cliente. Además es una metodología que requiere de pocas personas (básicamente 2, pero yo actuaré de forma virtual como 2 personas). XP se basa en la iteración frecuente a través de la cual los desarrolladores implementan historias de usuario. Las historias de usuario son declaraciones sencillas e informales del cliente sobre las funcionalidades necesarias³.

- Una historia de usuario es una descripción convencional del usuario sobre una característica del sistema requerido.
- No menciona detalles más finos como los diferentes escenarios que pueden ocurrir. Sobre la base de historias de usuario, el equipo del proyecto propone “Metáforas”. Las metáforas son una visión común de cómo funcionaría el sistema.
- El equipo de desarrollo puede decidir construir un “Spike” para alguna característica. Un Spike es un programa muy simple que se construye para explorar la idoneidad de una solución propuesta. Puede considerarse similar a un prototipo.

²<https://www.geeksforgeeks.org/software-engineering-extreme-programming-xp/>

³<http://www.jtech.ua.es/j2ee/2011-2012/restringido/met/sesion02-apuntes.html>

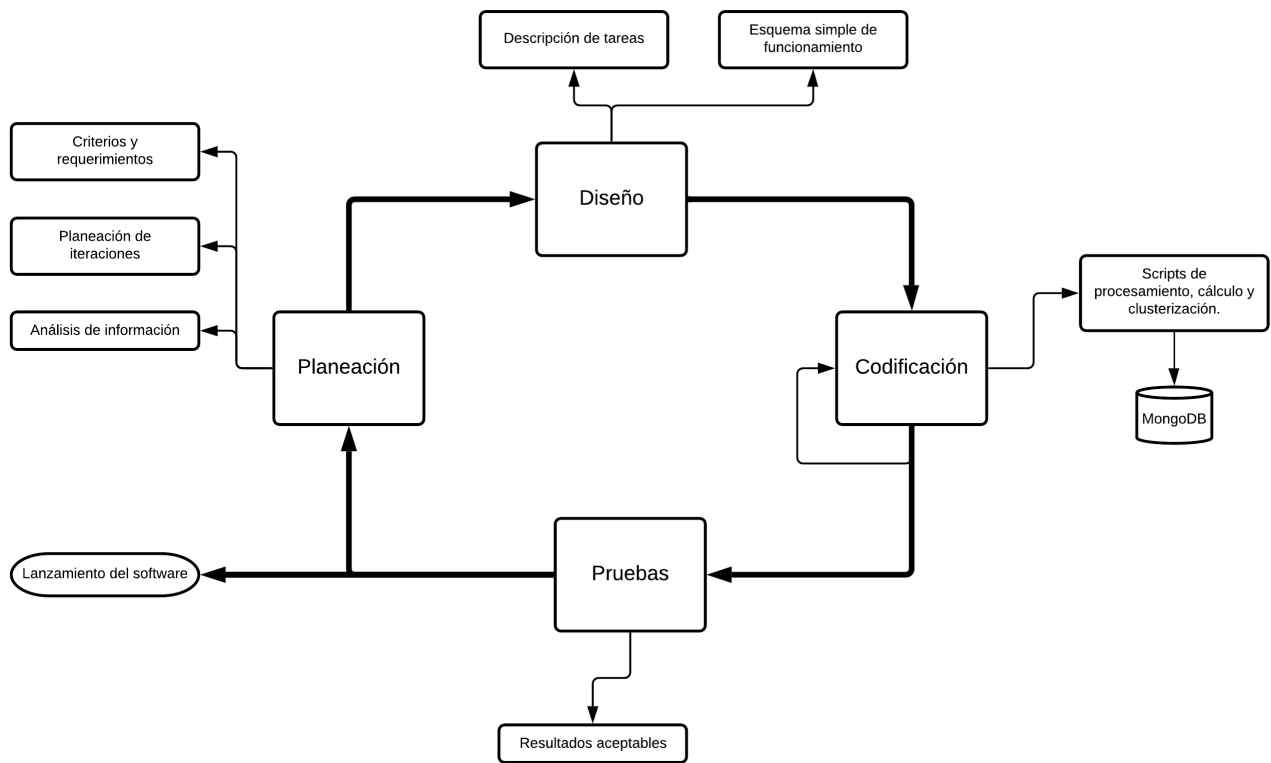


Figura 2: Representación de desarrollo del software utilizando la metodología XP.

En la Figura 2 se puede ver como primero se inicia con una planeación del proyecto, en el cual se debe hacer un análisis de la información, planear las iteraciones, anotar los criterios y requerimientos del sistema. Luego en el diseño se hará la descripción de tareas a desarrollar en dichas iteraciones utilizando un esquema simple de funcionamiento. Después de se realizara la codificación de los distintos scripts, ya sea de procesamiento de datos, de cálculo y de clusterización. Los cuales se guardarán en la base de datos no relacional MongoDB. Una vez terminada la codificación se realizarán pruebas, en las cuales se buscarán resultados relativamente aceptables para así poder lanzar el software a producción.

4. Desarrollo

4.1. Requerimientos

4.2. Diseño

El diseño general del sistema partirá con la utilización de los datos del Ministerio de Ciencia, que luego serán procesados para poder realizar los cálculos de las predicciones, luego esos resultados se almacenarán en una base de datos llamada MongoDB, después se ejecutarán los algoritmos de clusterización para hacer grupos de comunas y eso será mostrado en el servicio web.

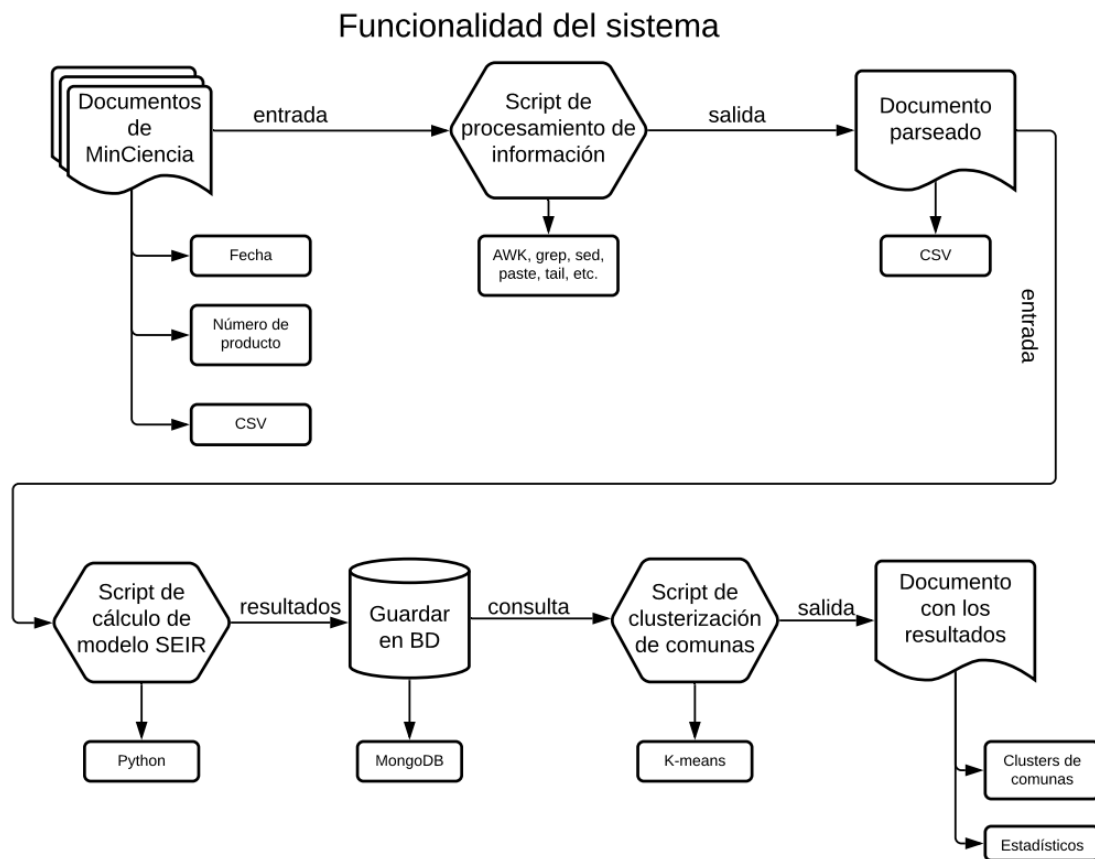


Figura 3: Funcionalidad general del sistema.

4.2.1. Base de datos

Se utilizará una base de datos no relacional llamada MongoDB, es una base de datos distribuida, basada en documentos y de uso general que ha sido diseñada para desarrolladores de aplicaciones modernas y para la era de la nube. Es una base de datos documental, lo que significa que almacena datos en forma de documentos tipo JSON⁴.

Las razones del porqué no se utilizará una base de datos relacional como PostgreSQL⁵, es porque por un lado se quiere utilizar lo aprendido en el módulo anterior y por que el servicio web solo será una visualización de información, por lo tanto las consultas que se harán a la base de datos serán muy sencillas.

La forma en como se utilizará la BD.....

Hablar de los documentos que se haran (los cuales seran los mismos que seran utilizados para mostrar la info en la pagina despues)

4.2.2. Tecnologías

Hablar de Django, python3, mongodb, de las librerías usadas en python, de las de visualización de la data, que es tema pendiente aun.

4.3. Implementación

4.3.1. Iteración 1, descarga y procesamiento de información.

1. Revisar y descargar la información entregada por el Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (disponible en <https://github.com/MinCiencia/Datos-COVID19>).
2. En la cual se filtrarán desde el producto 2 los casos confirmados de cada comuna.
3. Desde el producto 38 para filtrar los fallecidos por comuna.
4. No hay información de recuperados, así que se tendrá que calcular por fecha de inicio de los síntomas.

⁴<https://www.mongodb.com/es>

⁵<https://www.postgresql.org>

5. Luego de descargar los datos de casos confirmados, fallecidos y el documento con las fechas de inicio y término de síntomas, se procede a no descargar toda la información de todas las fechas (debido a su baja o nula consistencia), por lo cual se tomará la información más actualizada (solo lo de las últimas 2 semanas).
6. Se extraerá información de la migración de cada comuna desde el censo 2017 (<https://www.censo2017.cl>).
7. Programar código que hará el procesamiento de los datos para que quede un archivo csv con las siguientes columnas: nombre comuna, región a la que pertenece la comuna, población, contagiados, recuperados, fallecidos, tasa de migración entre comunas.
8. Se aplaza 1 semana mas, debido a problemas técnicos.

4.3.2. Iteración 2, programación de scripts para calcular el modelo SEIR y guardar sus resultados.

1. Este script se codificará con el lenguaje de programación Python (versión 3.6) y la utilización de las librerías Numpy, Scipy y Mongo.
2. Como input será un archivo en formato csv, con las columnas nombre comuna, región a la que pertenece la comuna, población, contagiados, recuperados, fallecidos, tasa de migración entre comunas.
3. Teniendo la información, se realizará el cálculo del modelo SEIR.
4. Guardar esta información resultante post-cálculo en la base de datos no relacional MongoDB.
5. Se aplaza 1 semana, por problemas técnicos.
6. Atraso, se aplaza 3 días mas.

5. Conclusiones

Referencias

- [1] Organización Mundial de la Salud
<https://www.who.int/es/health-topics/coronavirus>