



*Diagnóstico de
Enfermedades con
Inteligencia Artificial*

Por Michelle Díaz

@Michdiaz_

GESTIÓN DE REGISTROS/EXPEDIENTES MÉDICOS

ANÁLISIS DE ESTUDIOS/TESTS MÉDICOS

ANÁLISIS DE RAYOS X

CONSULTA DIGITAL

DISEÑO DE TRATAMIENTO

APLICACIONES

CUIDADO Y MONITOREO CONTINUO

MEDICINA GENÉTICA Y GENÓMANA

CREACIÓN DE MEDICAMENTOS

ANÁLISIS DE TOMOGRAFÍAS

GESTIÓN DE MEDICAMENTO

GESTIÓN DE REGISTROS/EXPEDIENTES MÉDICOS

ANÁLISIS DE ESTUDIOS/TESTS MÉDICOS

ANÁLISIS DE RAYOS X

CONSULTA DIGITAL

DISEÑO DE TRATAMIENTO

APLICACIONES

CUIDADO Y MONITOREO CONTINUO

MEDICINA GENÉTICA Y GENÓMANA

CREACIÓN DE MEDICAMENTOS

ANÁLISIS DE TOMOGRAFÍAS

GESTIÓN DE MEDICAMENTO

DIGITALIZACIÓN EM

Medical Records Release Form

Patient Name: _____ Date of Birth: _____

Person Requesting records and relationship: _____

Home Phone: _____ Daytime Phone: _____

By signing this form, I authorize you to release confidential health information about me, by releasing a copy of my medical records, or a summary or narrative of my protected health information, to the person(s) or entity listed below.

HIV/AIDS: I DO DO NOT consent to the release of any positive or negative test result for AIDS or HIV infection, antibodies to AIDS or infection with any other causative agent of AIDS with the rest of my medical records. Initial: _____ Date: _____

Limitations on the information you may release subject to this Release Form are as follows:

Release my protected health information to the following person(s)/entity:

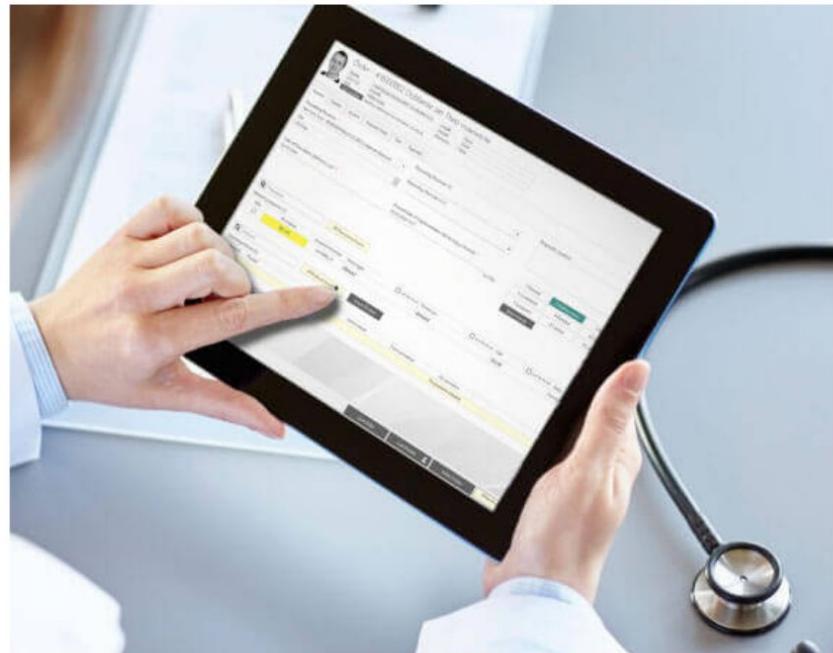
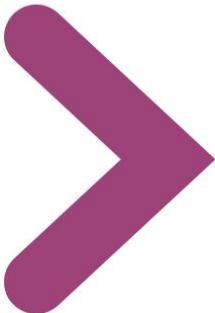
Name: _____
Street: _____
City: _____ State: _____ Zip: _____

I do do NOT give permission for these records to be faxed to the above entity.

The reasons or purposes for this release of information are as follows:

Patient Signature (or parent, guardian or legal representative) _____ Date _____

I understand that you will provide this information within 15 days from receipt of request and that a fee for preparing and furnishing this information may be charged according to rulings set forth by the Texas State Board of Medical Examiners.



GESTIÓN DE REGISTROS/EXPEDIENTES MÉDICOS

ANÁLISIS DE ESTUDIOS/TESTS MÉDICOS

ANÁLISIS DE RAYOS X

CONSULTA DIGITAL

DISEÑO DE TRATAMIENTO

APLICACIONES

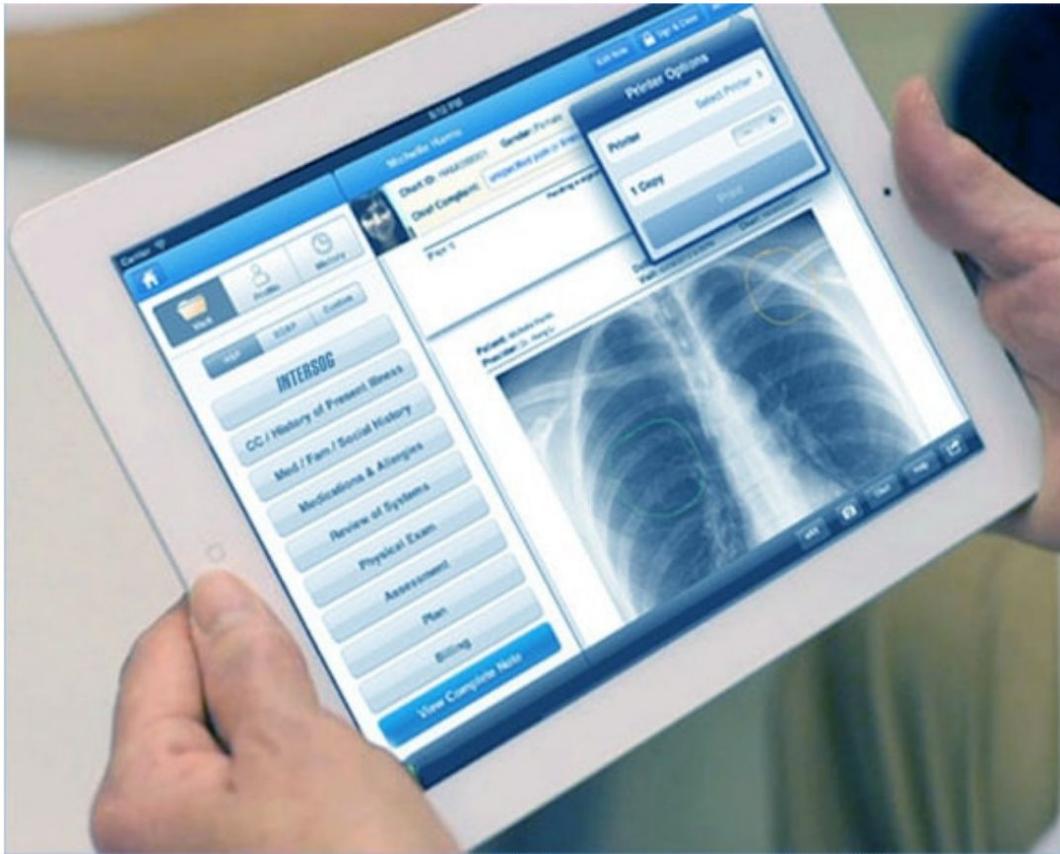
CUIDADO Y MONITOREO CONTINUO

MEDICINA GENÉTICA Y GENÓMANA

CREACIÓN DE MEDICAMENTOS

ANÁLISIS DE TOMOGRAFÍAS

GESTIÓN DE MEDICAMENTO



CheXNet:

Detección de neumonía a nivel de radiólogo en radiografías de tórax con Deep Learning.

<https://stanfordmlgroup.github.io/projects/chexnet/>

GESTIÓN DE REGISTROS/EXPEDIENTES MÉDICOS

ANÁLISIS DE ESTUDIOS/TESTS MÉDICOS

ANÁLISIS DE RAYOS X

CONSULTA DIGITAL

DISEÑO DE TRATAMIENTO

APLICACIONES

CUIDADO Y MONITOREO CONTINUO

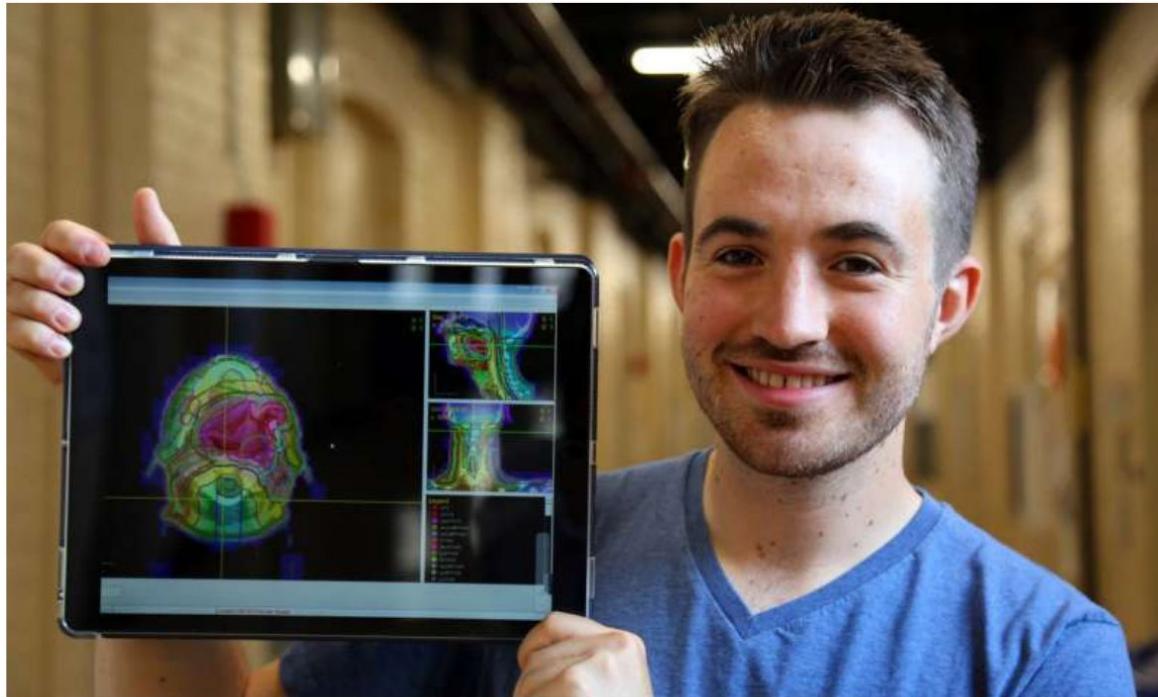
MEDICINA GENÉTICA Y GENÓMANA

CREACIÓN DE MEDICAMENTOS

ANÁLISIS DE TOMOGRAFÍAS

GESTIÓN DE MEDICAMENTO

Planificación automatizada basada en conocimiento para el cáncer de orofaringe.



<https://aapm.onlinelibrary.wiley.com/doi/full/10.1002/mp.12930>
<https://deepmind.com/applied/deepmind-health/>

GESTIÓN DE REGISTROS/EXPEDIENTES MÉDICOS

ANÁLISIS DE ESTUDIOS/TESTS MÉDICOS

ANÁLISIS DE RAYOS X

CONSULTA DIGITAL

DISEÑO DE TRATAMIENTO

APLICACIONES

CUIDADO Y MONITOREO CONTINUO

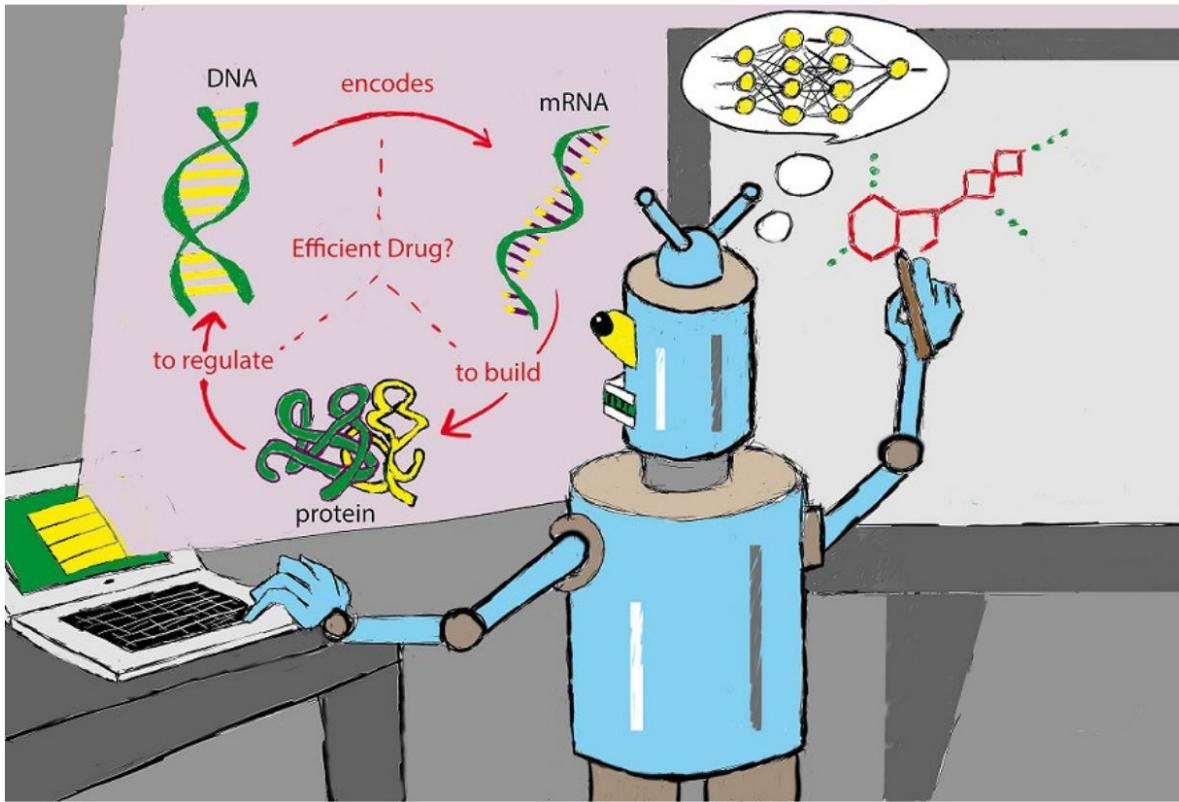
MEDICINA GENÉTICA Y GENÓMANA

CREACIÓN DE MEDICAMENTOS

ANÁLISIS DE TOMOGRAFÍAS

GESTIÓN DE MEDICAMENTO

Creación de Medicamentos y Medicina Genética y genómana



<https://www.nature.com/articles/d41586-018-05267-x>

GESTIÓN DE REGISTROS/EXPEDIENTES MÉDICOS

ANÁLISIS DE ESTUDIOS/TESTS MÉDICOS

ANÁLISIS DE RAYOS X

CONSULTA DIGITAL

DISEÑO DE TRATAMIENTO

APLICACIONES

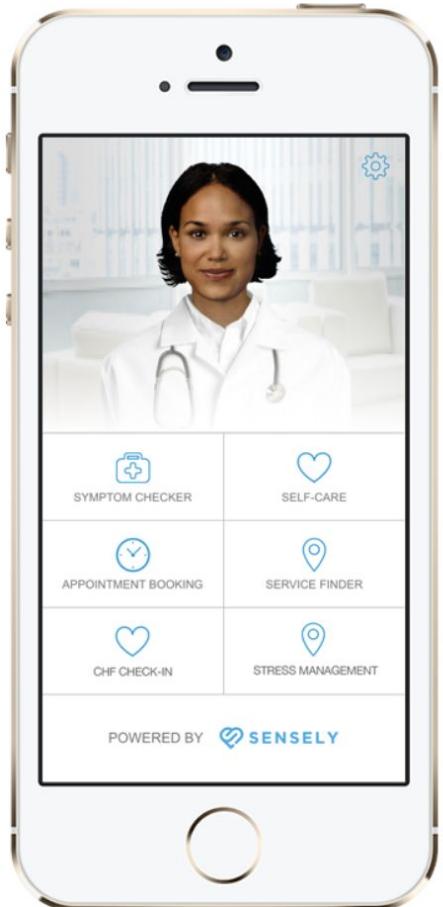
CUIDADO Y MONITOREO CONTINUO

MEDICINA GENÉTICA Y GENÓMANA

CREACIÓN DE MEDICAMENTOS

ANÁLISIS DE TOMOGRAFÍAS

GESTIÓN DE MEDICAMENTO



Análisis de Estudios

<https://arxiv.org/abs/1808.04928>

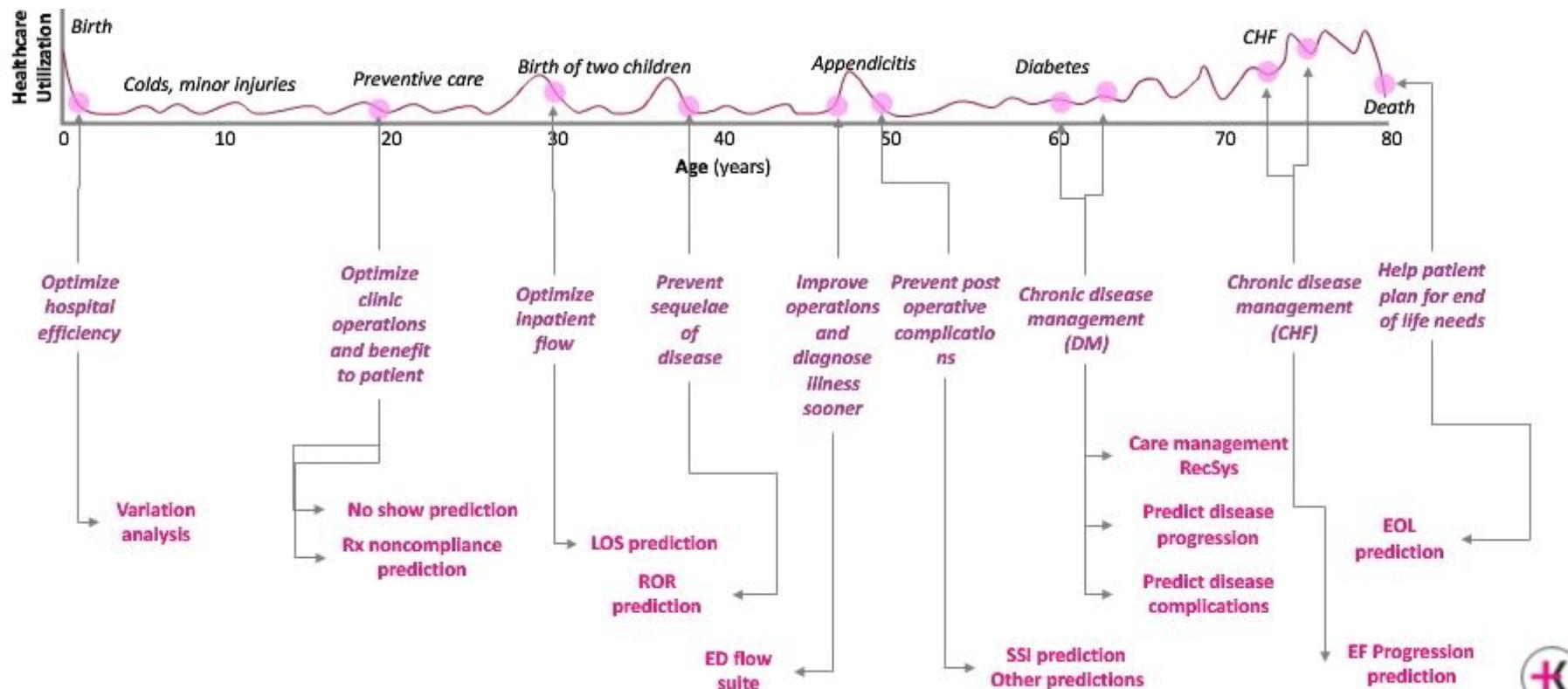
Cuidado y Monitoreo Continuo

<https://www.wellframe.com/>

Gestión de Medicamento

<https://news.stanford.edu/2018/07/10/ai-predicts-drug-pair-side-effects/>

ML puede mejorar el servicio médico a través del cuidado continuo.



¿Cómo son tomadas las decisiones actualmente?



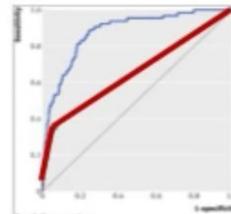
1. Heuristics



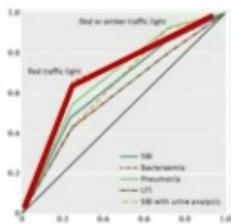
2. Rules based system



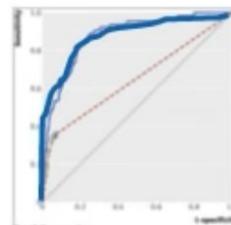
3. ML based System



~80% Care Decisions



~18% Care Decisions



~2% Care Decisions

Caso de Estudio

Disease Prediction by Machine Learning Over Big Data From Healthcare Communities

MIN CHEN¹ , (Senior Member, IEEE), YIXUE HAO¹ , KAI HWANG² , (Life Fellow, IEEE), LU WANG¹ , AND LIN WANG^{3,4} 1School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China 2University of Southern California, Los Angeles, CA 90089, USA 3Research Center for Tissue Engineering and Regenerative Medicine, Huazhong University of Science and Technology, Wuhan 430022, China 4Department of Clinical Laboratory, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China. Corresponding author: Lin Wang (lin_wang@hust.edu.cn)

¿Por qué es interesante el proyecto?

- Datos en otro lenguaje

¿Por qué es interesante el proyecto?

- Datos en otro lenguaje
- Datos estructurados

¿Por qué es interesante el proyecto?

- Datos en otro lenguaje
- Datos estructurados
- Datos NO estructurados

¿Por qué es interesante el proyecto?

- Datos en otro lenguaje
- Datos estructurados
- Datos NO estructurados
- Múltiples algoritmos

¿Con qué recursos contaban?

- Datos reales de 31,919 pacientes con 20,320,848 registros en total
- Información Estadística
- Resultados de Exámenes médicos
- Historial de Enfermedades
- Expedientes electrónicos
- Datos de imágenes médicas
- Datos genéticos

Datos Específicos

Datos Estructurados

- Datos de laboratorio
- Información básica del paciente (edad, género, hábitos de vida, entre otros).

Datos No Estructurados

- Narración del paciente de su enfermedad
- Anamnesis
- Resultado del diagnóstico del médico

Data category	Item	Description
Structured data	Demographics of the patient	Patient's gender, age, height, weight, etc.
	Living habits	Whether the patient smokes, has a genetic history, etc.
	Examination items and results	Includes 682 items, such as blood, etc.
	Diseases	Patient's disease, such as cerebral infarction, etc.
Unstructured text data	Patient's readme illness	Patient's readme illness and medical history
	Doctor's records	Doctor's interrogation records

Limitantes

- Chino
- Diferencia de Enfermedades en diferentes regiones
 - Climas
 - Hábitos

Desafíos

- ¿Cómo deben abordarse los datos faltantes?
- ¿Cómo deben determinarse las principales enfermedades crónicas en una región determinada y las principales características de la enfermedad en la región?
- ¿Cómo se puede usar la tecnología de análisis de datos grandes para analizar la enfermedad y crear un mejor modelo?

Análisis y Descubrimientos

Obtuvieron estadísticas sobre el número de pacientes, la proporción de sexos de los pacientes y la enfermedad principal en determinada región (China Central), cada año a partir de los datos de texto estructurados y no estructurados, los resultados estadísticos son:

- La proporción de pacientes masculinos y femeninos hospitalizados cada año tiene poca diferencia
- Hay más pacientes ingresados en el hospital en 2014.
- La hospitalización por enfermedades crónicas siempre ha estado ocupando una gran proporción en esta área a través de las estadísticas de los datos.
- El número de pacientes hospitalizados con enfermedades crónicas de infarto cerebral, hipertensión y diabetes representó el 5.63% del número total de pacientes ingresados en el hospital en 2015, mientras que las otras enfermedades ocuparon una pequeña proporción.

Statistics	2013	2014	2015
Number of inpatients	7265	24756	10552
Males	42.88%	50.36%	57.60%
Females	57.12%	49.64%	42.40%
Proportion of patients with cerebral infarction	1.47%	1.01%	1.66%
Proportion of hypertensive patients	1.06%	1.04%	1.98%
Proportion of diabetics	1.17%	0.99%	1.99%

Predictión

Aplican el modelo de predicción de riesgo para el infarto cerebral como los métodos de aprendizaje supervisado de ML, es decir, el valor de entrada es el valor de los atributos del paciente.

$$X = (x_1, x_2, \dots, x_n)$$

Predicción

Aplican el modelo de predicción de riesgo para el infarto cerebral como los métodos de aprendizaje supervisado de ML, es decir, el valor de entrada es el valor de los atributos del paciente.

$$X = (x_1, x_2, \dots, x_n)$$

Edad, sexo, prevalencia de los síntomas y los hábitos de vida (fumar o no) y otros datos estructurados y datos no estructurados.

Predicción

El valor de salida es C, que indica si el paciente se encuentra entre la población de alto riesgo de infarto cerebral.

$$C = \{C_0, C_1\}$$

Dónde:

C₀ indica que el paciente está en alto riesgo de infarto cerebral,
C₁ indica que el paciente está en bajo riesgo de infarto cerebral.

Decisiones

- Combinar los datos estructurados y no estructurados en el campo de la salud para evaluar el riesgo de enfermedad.

Decisiones

- Combinar los datos estructurados y no estructurados en el campo de la salud para evaluar el riesgo de enfermedad.
- Utilizar el modelo de variable latente para reconstruir los datos faltantes de los registros médicos recopilados del hospital.

Decisiones

- Combinar los datos estructurados y no estructurados en el campo de la salud para evaluar el riesgo de enfermedad.
- Utilizar el modelo de variable latente para reconstruir los datos faltantes de los registros médicos recopilados del hospital.
- Utilizar el conocimiento estadístico, para determinar las principales enfermedades crónicas en la región.

Decisiones

- Combinar los datos estructurados y no estructurados en el campo de la salud para evaluar el riesgo de enfermedad.
- Utilizar el modelo de variable latente para reconstruir los datos faltantes de los registros médicos recopilados del hospital.
- Utilizar el conocimiento estadístico, para determinar las principales enfermedades crónicas en la región.
- Para manejar datos estructurados, consultaron con expertos del hospital para extraer features útiles. Para datos de texto no estructurados, seleccionaron los features automáticamente usando CNN.

Decisiones

- Proponen un nuevo algoritmo de predicción de riesgo de enfermedad multimodal (CNN-MDRP) basado en CNN para datos estructurados y no estructurados. El modelo de riesgo de enfermedad se obtiene mediante la combinación de características estructuradas y no estructuradas.

Decisiones de configuración

- Selección de 706 pacientes en total como los datos del experimento
- División de los datos aleatoriamente en datos de entrenamiento y de prueba.

La proporción entre el dataset de entrenamiento y el dataset de prueba es 6: 1, es decir, 606 pacientes como conjunto de datos de entrenamiento, mientras que 100 pacientes como conjunto de datos de prueba.

Decisiones de desarrollo

Para los datos Estructurados para predecir el riesgo de infarto cerebral:

- Naive Bayesian (NB)
- K-Neighbor más cercano (KNN)
- Decision Tree (DT) algorítmicos

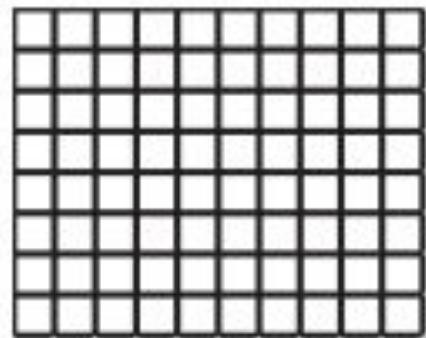
Decisiones de desarrollo

Para los datos de Texto

Algoritmo de predicción del riesgo de enfermedad unimodal (CNN-UDRP) basado en CNN para predecir el riesgo de enfermedad infarto cerebral.

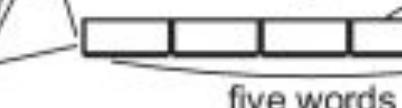
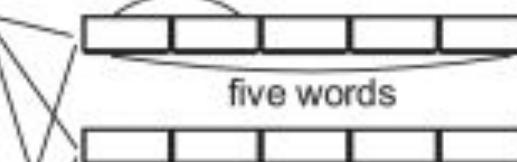
convolution filters : 100
region size : 50×5

50×5



word embedding size : 50

addPadding



(b)

convolution



convolution

max
pooling

full
connection

(d) (e)

(g)
feature
Fusion



(d) (e)

output

(f)



structure data feature : 79

Complejidad

- Datos Sintácticos**

Formato

Sexo = M/F =1/0

- Datos morfológicos**

Presión Sanguínea de 500

- Datos Semánticos**

Niños v.s. Adultos



Doctor, are you
sure that I am
ok now? My
head still hurts.

Well, the computer
assures me that
you are fine and it
has given me 100
reasons for it.

¡GRACIAS!