

Solicitud de Registro de Tema para Titulación

Quien suscribe, estudiante regular de la Maestría en Ciencia de Datos e Información

Nombre:	FRANCISCO EDUARDO ZAVALA RODRIGUEZ
----------------	---

Solicitamos ante el Coordinador Académico de la Maestría en Ciencia de Datos e Información, la autorización y registro para titulación en la Maestría, en la modalidad:

	Reporte Analítico de Experiencia Laboral
	Propuesta de Intervención
x	Implementación de un Proyecto
	Solución Estratégica

TITULO TENTATIVO DEL PROYECTO	Reconocimiento de texto para clasificación de causas básicas CIE-10
--	---

Solicitud de Registro de Tema para Titulación

Introducción.

El ISSSTE es una dependencia gubernamental que brinda no únicamente servicios para pensiones para trabajadores de gobierno si no que también una de sus principales funciones es brindar un servicio de calidad a los derechohabientes que acuden a sus unidades médicas las cuales están divididas en tres grandes grupos:

1. Unidades de primer nivel, dentro de las cuales se encuentran UMF (unidades de medicina familiar), CMCT (centros médicos en centros de trabajo), CMF (clínicas de medicina familiar) y CAF (centros de atención familiar), por mencionar algunas dentro de estas unidades, una de sus principales funciones es detectar y prevenir a futuro enfermedades crónicas degenerativas así como también atender padecimientos comunes como lo son gripas, problemas de nutrición, planificación familiar, caries, etc.
2. Unidades de segundo nivel, como son HG (hospitales generales), CH (clínicas hospital), CMFEQ (clínicas de medicina familiar y quirófano) y CE (clínicas de especialidad) en estas unidades se atienden a pacientes que no pudieron ser controlados en las unidades de primer nivel o que requieren alguna intervención quirúrgica simple, fracturas, urgencias menores o alguna atención especial de algún padecimiento presentado por el derechohabiente.
3. Unidades de tercer nivel aquí se encuentran HR Y CMN (centro médico nacional 20 de noviembre), estas unidades únicamente atienden padecimientos crónicos degenerativos o pacientes con problemas graves de salud, como puede ser cáncer, diabetes mellitus en fase terminal, problemas cardíacos, trasplantes, casos no reconocidos o de estudio de alguna enfermedad.

Todas estas unidades generan información diaria sobre padecimientos, muertes, estudios, consultas médicas, accidentes, personal que labora, etc. y si esta información no tiene un buen tratamiento puede ocasionar problemas como desabasto de medicamento, la mala adquisición de equipo médico, mala contratación de personal médico (enfermeras, médicos especialistas paramédicos, personal administrativo, etc.).

Por ende, la información y un buen análisis de estos datos que se generan ayudan a la institución a poder hacer cada día mejor su trabajo.

Antecedentes

La problemática surge de la revisión y validación de médica enfocada a muertes maternas, fetales, defunciones (padecimiento por el cual una persona muere) y morbilidades (padecimiento o causa de ingreso de un paciente a un hospital o unidad médica), sin embargo el mal diagnóstico elaborado por un sistema llamado "IRIS" ya que este se basa en el diagnóstico del médico capturado en el certificado de defunción y según las causas básicas asignadas en él, sin embargo al hacer la revisión

Solicitud de Registro de Tema para Titulación

de estos datos, resulta que el diagnóstico realizado por el médico y la causa básica de muerte no coincide con la dictada por el programa ya antes mencionado.

Esto ocasiona que la confiabilidad del sistema a realizar codificaciones por medio de la CIE-10 no tenga mucha validez y se tenga que revisar dato por dato comparando el diagnóstico del médico contra el del sistema, ya que, si esto no se revisa adecuadamente, al presentar la información sin un proceso previo de validación, tanto las estadísticas que se alimentan de estos datos, así como indicadores serían erróneos.

Objetivo General

El objetivo general de este trabajo de tesis es encontrar el mejor clasificador para la búsqueda y coincidencia de los padecimientos presentados por los derechohabientes dentro de las unidades de segundo y tercer nivel, donde los principales errores se cometen al hacer la captura de la información registrada dentro de los certificados de defunción y dentro del sistema interno del instituto.

Para esto nos apoyaremos en el sistema internacional IRIS desarrollado en Alemania en 2007, el cual realiza las clasificaciones de morbilidades y mortalidad por medio de catálogos registrados en su sistema y por medio de algoritmos de inteligencia artificial para realizar el mejor diagnóstico.

Objetivos específicos.

Además, con ello poder observar y comparar los resultados obtenidos con el algoritmo elaborado de manera interna y poder tomar una muestra para comparar los resultados obtenidos de la realización de las clasificaciones de mortalidad y morbilidad y poder definir cuál de los dos algoritmos es mejor que el otro.

Resultados Esperados

Como uno de los principales resultados a los que se pretende llegar es que el clasificador elaborado tenga mejor tasa de aciertos sobre la clasificación de enfermedades y muertes que se registran dentro del instituto.

Solicitud de Registro de Tema para Titulación

índice planteado para el tema de tesis

- 1 Importancia de la codificación CIE-10
 - 1.1 Cie-10
 - 1.2 IRIS
 - 1.3 Causas básicas de defunción
- 2 Procesamiento de lenguaje natural
 - 2.1 ¿Qué es el procesamiento de lenguaje natural?
 - 2.2 Aplicaciones de NLP
- 3 Maquinas de Soporte vectorial en la clasificación de texto
 - 3.1 Modelos de pesado
 - 3.2 Modelos se similitud
- 4 Word embedding o incrustación de palabras
- 5 Resultados obtenidos
- 6 Conlusiones
- 7 Bibliografía

Solicitud de Registro de Tema para Titulación

- <https://www.utm.mx/~jahdezp/archivos%20estructuras/DESICION.pdf>
- https://ccc.inaoep.mx/~esucar/Clases-mgp/Proyectos/MGP_RepProy_Abr_29.pdf
- <https://zaquan.unizar.es/record/59156/files/TAZ-TFG-2016-2057.pdf>
- <https://medium.com/datos-y-ciencia/algoritmos-naive-bayes-fudamentos-e-implementaci%C3%B3n-4bcb24b307f>
- https://sistemas.fciencias.unam.mx/~erhc/calculo3_20171/derivadas_parciales_dir_eccionales_2016_12.pdf
- <https://medium.com/datos-y-ciencia/aprendizaje-no-supervisado-en-machine-learning-agrupaci%C3%B3n-bb8f25813edc>
- https://sisbib.unmsm.edu.pe/bibvirtualdata/Tesis/Basic/Salcedo_pc/enPDF/Cap2.PDF
- https://www.cienciadedatos.net/documentos/35_principal_component_analysis
- <https://ais.paho.org/classifications/chapters/pdf/volume1.pdf>
- https://rccs.cic.ipn.mx/2016_128/RENTOL_%20Un%20algoritmo%20de%20agrupamiento%20basado%20en%20K-means.pdf
- https://ccc.inaoep.mx/~esucar/Clases-mgp/Proyectos/MGP_RepProy_Abr_29.pdf
- <https://powerbi.microsoft.com/en-us/>