



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
CARRERA DE SOFTWARE**

**TRABAJO DE INTEGRACIÓN CURRICULAR ACADÉMICA PREVIA A LA OBTENCIÓN
DEL TÍTULO
DE INGENIERO EN SOFTWARE**

**TEMA: SISTEMA DE PREDICCIÓN DE DIAGNÓSTICOS MÉDICOS EN BASE A LAS NOTAS
CLÍNICAS DE LOS PACIENTES, APLICANDO TÉCNICAS Y MODELOS DE APRENDIZAJE
AUTOMÁTICO.**

AUTORES:

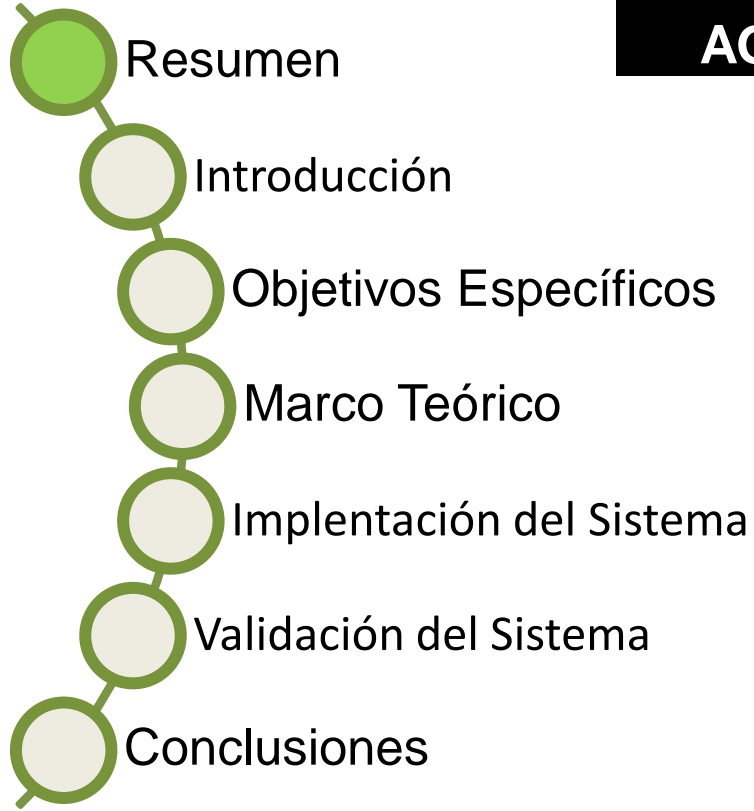
**CRUZ CAIZA, MACARENA LIZBETH
QUISHPE ROCHA, LUIS LENIN**

DIRECTOR:

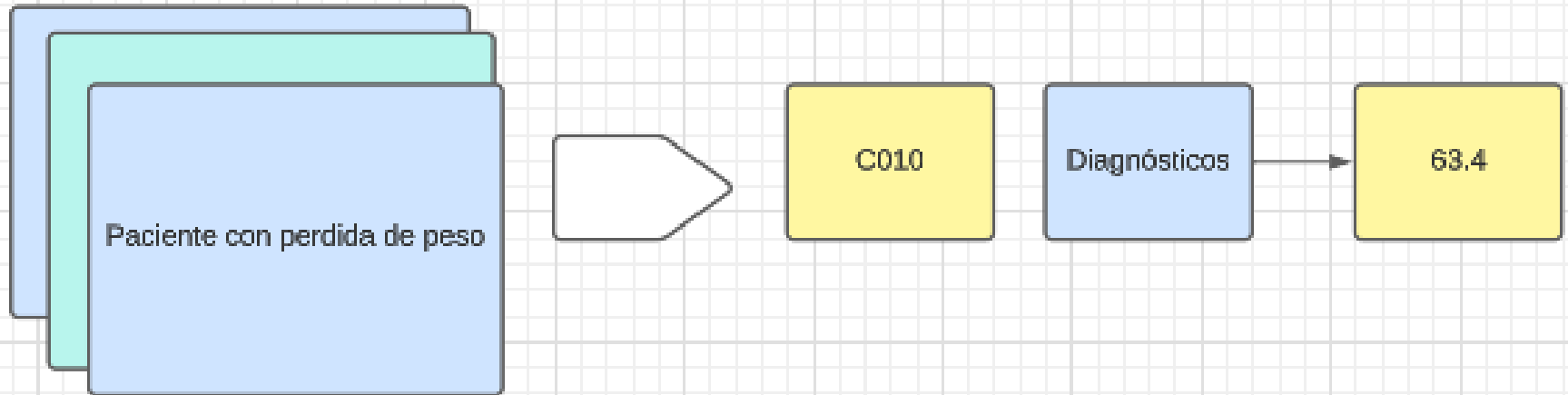
**ING. UYAGUARI UYAGUARI, ALVARO DANILO
LATACUNGA FEBRERO, 2023**



AGENDA



PLANTEAMIENTO DEL PROBLEMA



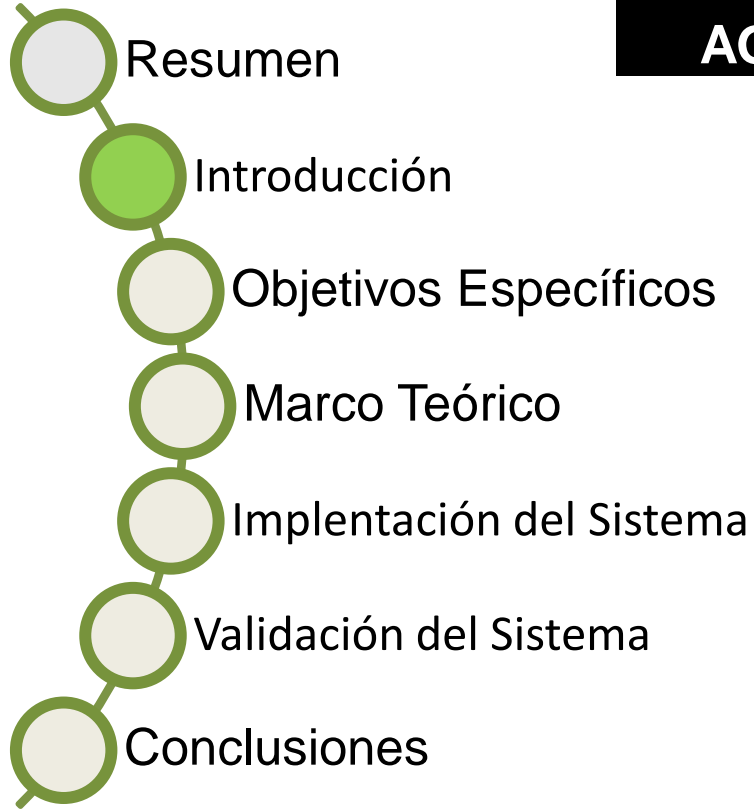
RESUMEN

SISTEMA DE PREDICCIÓN DE DIAGNÓSTICOS MÉDICOS EN BASE A LAS NOTAS CLÍNICAS DE LOS PACIENTES

En el presente trabajo se desarrolló un sistema de predicción de diagnósticos médicos en base a notas clínicas de los pacientes, aplicando técnicas y modelos de aprendizaje automático, con el objetivo de precisar la toma de decisiones mediante técnicas de análisis de datos. Los modelos de predicción clínica se aplican comúnmente en la práctica médica para ayudar a los profesionales de la salud a determinar el diagnóstico o pronóstico de un paciente.



AGENDA



INTRODUCCIÓN

PROPÓSITO Y CONTEXTUALIZACIÓN DEL TEMA

El propósito es desarrollar un sistema y nuevos métodos para la identificación de entidades biomédicas en español utilizando técnicas de procesamiento de lenguaje natural. Estos nuevos métodos y algoritmos serán integrados a un sistema para la predicción de diagnósticos, aplicando herramientas, métodos y buenas prácticas de ingeniería de software.



INTRODUCCIÓN

JUSTIFICACIÓN DEL INTERÉS DE LA INVESTIGACIÓN

El procesamiento natural del lenguaje (PNL) es el campo de estudio que busca entender cómo funciona el lenguaje. Entre estas tareas que maneja se encuentra la generación de un nuevo texto, traducciones de un idioma a otro, preguntas y respuestas, generar resumen, chatbots entre otros. Los modelos predictivos son un grupo de técnicas que mediante los campos del aprendizaje automático la recolección de datos y el reconocimiento de patrones, pretende dar una predicción de resultados futuros.



AGENDA

- Resumen
- Introducción
- Objetivos
- Marco Teórico
- Implentación del Sistema
- Validación del Sistema
- Conclusiones



OBJETIVOS

Desarrollar un sistema de predicción de diagnósticos médicos en lenguaje español.

Explorar modelos y corpus para el proceso de predicción de diagnósticos médicos.

Desarrollar un modelo de predicción de diagnósticos en base a textos médicos en lenguaje español.

Aplicar buenas prácticas en el ciclo de desarrollo e implementar el sistema con frameworks y arquitecturas actuales.

Redacción de la tesina.



AGENDA

- Resumen
- Introducción
- Objetivos Específicos
- Marco Teórico
- Implentación del Sistema
- Validación del Sistema
- Conclusiones



MARCO TEÓRICO

Procesamiento de lenguaje natural

PLN consiste en utilizar el lenguaje natural para comunicarse con una computadora, la computadora tiene que entender las oraciones provistas, estos lenguajes naturales facilitan el desarrollo de programas que realizan tareas relacionadas con el lenguaje y el desarrollo de modelos que nos ayudan a comprender los mecanismos humanos asociados al lenguaje.



MARCO TEÓRICO

Procesamiento de lenguaje natural aplicado en el campo médico

El propósito de codificar texto clínico es estructurar la información de tal manera que pueda ser fácilmente utilizada para tareas de gestión, clasificación de enfermedades, estadísticas o toma de decisiones.

Ejemplos de codificación son la Clasificación Internacional de Enfermedades (CIE) y la Nomenclatura Sistematizada de Términos Médicos Clínicos (SNOMED-CT)



AGENDA

- Resumen
- Introducción
- Objetivos Específicos
- Marco Teórico
- Implentación del Sistema
- Validación del Sistema
- Conclusiones



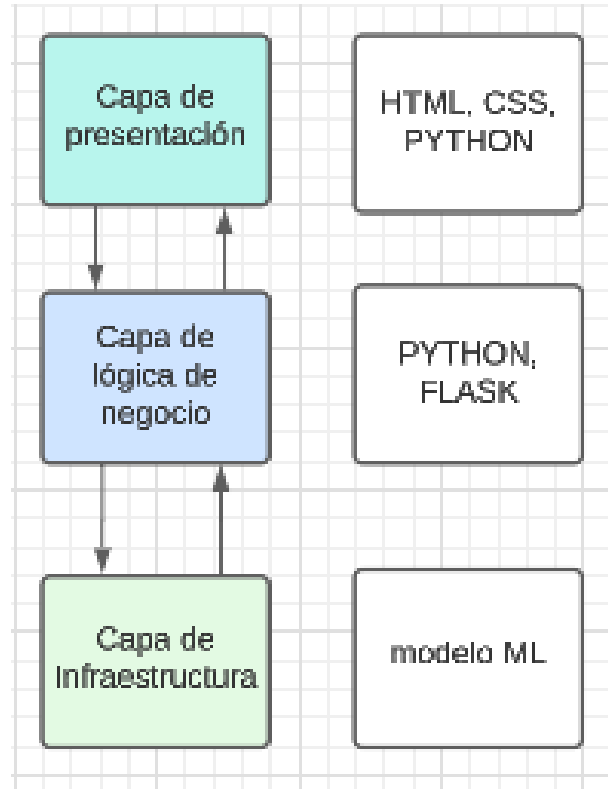
IMPLEMENTACIÓN DEL SISTEMA

Corpus usado para el sistema

Para el sistema se usa un corpus médico, el cual contiene un conjunto de datos que son los casos clínicos, en el cual se encuentran datos de entrenamiento, desarrollo y pruebas, todas las anotaciones de estas notas clínicas se encuentran en “CodiEsp corpus: gold estándar Spanish clinical cases coded in ICD10 (CIE10)-eHealth CLEF2020”.



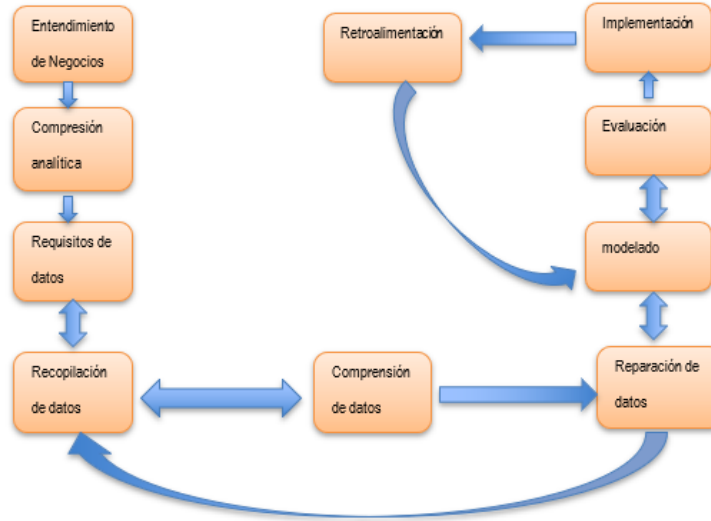
ARQUITECTURA DEL SISTEMA



IMPLEMENTACIÓN DEL SISTEMA

Metodología de ciencia de datos

Las metodologías usadas en ese ámbito, la ciencia de datos se considera un enfoque nuevo y prometedor utilizado en la recopilación y el análisis de datos en muchas disciplinas.



IMPLEMENTACIÓN DEL SISTEMA

Metodología de desarrollo ágil

En la metodología Scrum existen eventos que permiten el control de forma periódica para mostrar el avance en el desarrollo del proyecto evitando inconvenientes de tiempo.



Tomado, Drew, 2023, *Ventajas y desventajas de la metodología Scrum*. Wearedrew.co. <https://blog.wearedrew.co/productividad/-ventajas-y-desventajas-de-la-metodologia-scrum>



AGENDA

- Resumen
- Introducción
- Objetivos Específicos
- Marco Teórico
- Implentación del Sistema
- Validación del Sistema
- Conclusiones



VALIDACIÓN DEL SISTEMA

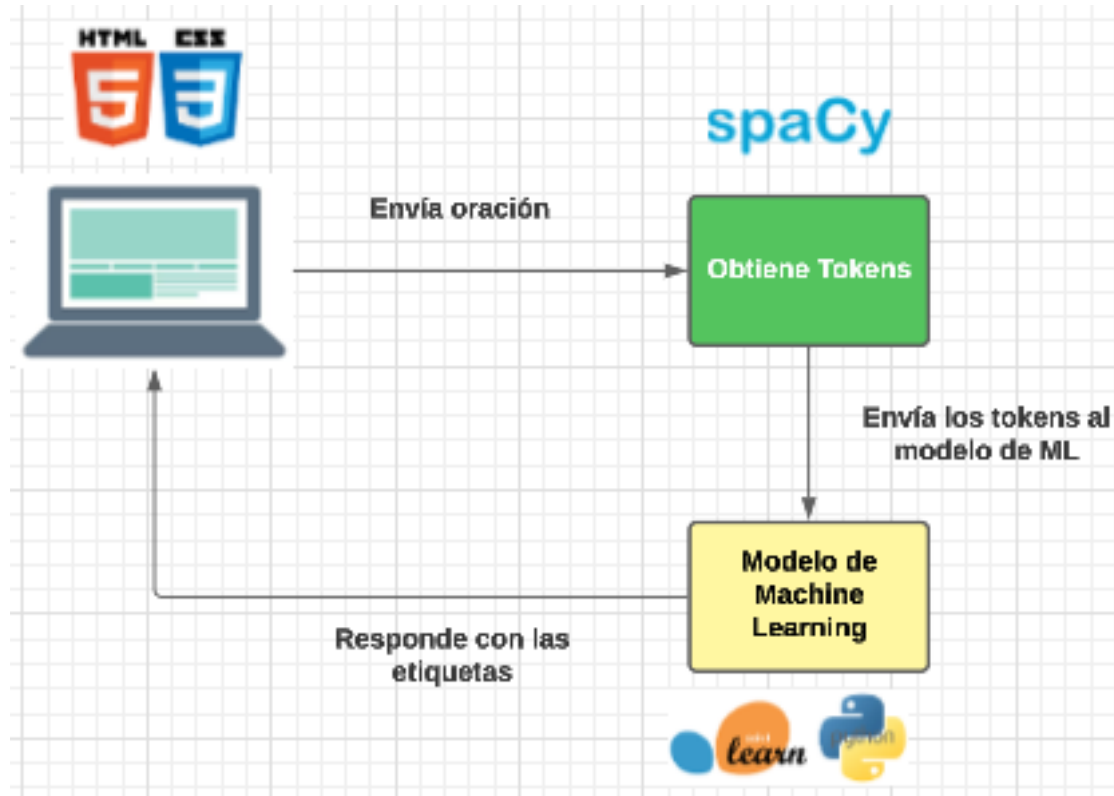
Métricas de Precisión

Profundizar las métricas que son la exactitud, precisión, la sensibilidad y la especificación.

Accuracy	Precisión	F1	Recall
0.9145102078704508	0.9218187124288042	0.8736740257232903	0.9145102078704508



SISTEMA WEB



SISTEMA WEB

Formulario de consulta

Entrada

Enviar texto

Se realizó una nueva ecografía en la que se aprecia una imagen compatible con quiste hidatídico de gran tamaño

Salida:

ecografía es B-PROCEDIMIENTO
quiste es B-DIAGNOSTICO
hidatídico es I-DIAGNOSTICO



AGENDA

- Resumen
- Introducción
- Objetivos Específicos
- Marco Teórico
- Implentación del Sistema
- Validación del Sistema
- Conclusiones



Conclusiones

Se cumplió con el objetivo de desarrollar un sistema de predicción de diagnósticos médicos en base a notas clínicas de los pacientes, aplicando técnicas y modelos de aprendizaje automático.

El desarrollo del marco teórico permitió la obtención de conocimientos acerca de los modelos de ML, técnicas y herramientas adecuadas para el desarrollo del sistema.

El uso de métricas de evaluación ayudó al análisis de resultados permitiendo validar la precisión del modelo y evidenciando si se encuentran las entidades de diagnósticos y procedimientos.

Al utilizar paperspace como entorno se consiguió un mejor rendimiento con respecto al tiempo de ejecución y así se logró entrenar el modelo con gran cantidad de datos.



Gracias



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA