



MAESTRÍA EN CIENCIA DE DATOS

Temas de tesis

Dr. José Eduardo Ochoa Luna



Universidad Católica
San Pablo

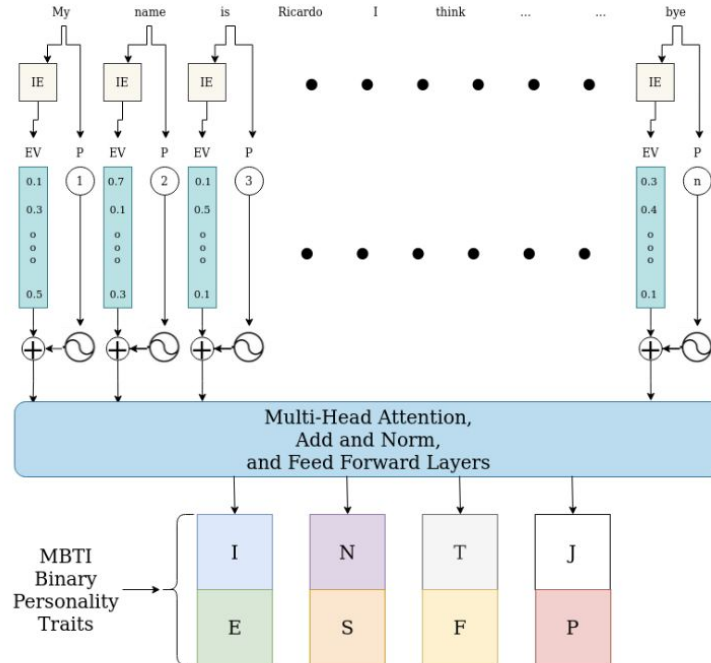
Departamento de Ciencia
de la Computación

12 Noviembre 2023

Personality Detection

MBTI Kaggle

- BERT Base.
- BERT Base Multilingual.
- DistilBERT Base.
- DistilBERT Base Multilingual.
- XLNet Base.
- RoBERTa Base.
- RoBERTa Distil.
- XLM-RoBERTa.
- ELECTRA Base.
- ELECTRA Small.



Suicidal Tendencies

1.4. La personalidad y el suicidio



NER epidemiología

- Se tiene un corpus anotado (NER) de enfermedades en LATAM
- Objetivo: proponer herramientas IA generativa para extraer información de ProMed-mail
- Ejemplo: Cuántos casos de sarampión hubo en Perú de marzo a abril del 2023?

➤ [BMC Bioinformatics](#). 2022 Dec 23;23(1):558. doi: 10.1186/s12859-022-05094-y.

Digital surveillance in Latin American diseases outbreaks: information extraction from a novel Spanish corpus

Antonella Dellanzo ¹, Viviana Cotik ² ³, Daniel Yunior Lozano Barriga ⁴, Jonathan Jimmy Mollapaza Apaza ⁴, Daniel Palomino ⁴, Fernando Schiaffino ⁵, Alexander Yanque Aliaga ⁴, José Ochoa-Luna ⁴

Affiliations + expand

PMID: 36564712 PMCID: [PMC9780622](#) DOI: [10.1186/s12859-022-05094-y](#)

[Free PMC article](#)

Marcadores basados en lenguaje natural para la discriminación de enfermedades neurodegenerativas

- **Proyecto Financiado por ProCiencia - 2023**
- Dr. Marcio Soto (Psicología), Investigador Principal, Universidad Católica San Pablo
- Dr. José Ochoa, Co-investigador, Universidad Católica San Pablo
- 1 Tesista de Maestría, Universidad Católica San Pablo
- 1 viaje a Buenos Aires para capacitación



LLMs con contexto LATAM

- Rafael Calvo: investigador principal (Imperial College)
- Proyecto aún en evaluación (Febrero 2024)
- Objetivo: proponer LLMs que no tengan bias, sean culturalmente representativos de culturas de LATAM.

Imperial College
London

Otros temas de interés

- Mantenimiento Predictivo con Deep Learning.
- Question Answering con IA Generativa.
- RAGs (Retrieval Augmented Generation).
- Herramientas de Ciencia de Datos asistidas por IA Generativa