

EconMentor

Integración de ChatBot con Tesis de la PUCP

Alexander Quispe¹ Rodrigo Grijalba²

¹The World Bank, PUCP - alexander.quispe@pucp.edu.pe

²PUCP, rgrijalba@pucp.edu.pe

Julio 2023

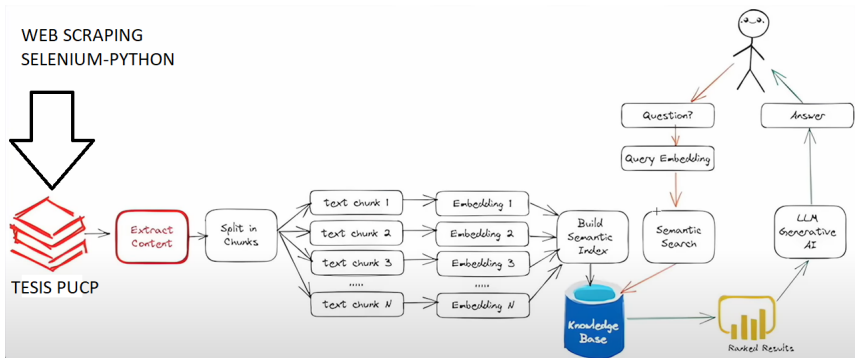
¿Qué es EconMentor?

- **ChatBot/Aplicación** que integra el modelo de **inteligencia artificial GPT-3.5** para atender consultas sobre **tesis de Economía PUCP**
- Facilita una **interacción Dinámica** entre el usuario y el repositorio de Tesis
- Dirigido a estudiantes de pregrado/Postgrado en búsqueda de temas de investigación para tesis
- Logra aconsejar sobre **nuevas preguntas de investigación en Economía**
- **Este ChatBot puede ser escalado a todas las especialidades**

¿Cómo funciona?

- Conformado por dos componentes:
 - ▶ **"Base de Conocimientos"**: Vectores que representan los contenidos de las tesis.
 - ▶ **Modelo de lenguaje**: Entrenado y provisto por OpenAI.
 - ▶ **Prompt Engineering**: Hemos instruido al robot a resolver consultas sobre tesis
- Ambos componentes son integrados por medio de algoritmos de búsqueda que hallan los documentos más relevantes en la Base de Conocimientos.
- El modelo responde a las consultas utilizando estos documentos

¿Cómo funciona?



¿Cómo se compara con el actual sistema?

Buscar

desigualdad de ingresos

Ir

Mostrar filtros avanzados

Mostrando ítems 1-10 de 56

0



Cambiar todo para que nada cambie: la independencia y la **desigualdad** económica entre 1792 y 1827

Camacho Gavilán, Abel Fernando (Pontificia Universidad Católica del Perú, 2020-07-04)

El presente trabajo estima el nivel de **desigualdad en ingresos** para Perú en el periodo comprendido entre 1792 y 1827. Para ello se constroyen tablas sociales a partir de los censos de los años mencionados y múltiples fuentes complementarias. Las...



Determinantes de la **desigualdad de ingresos** en Perú entre 2012 y 2020: un análisis a nivel de dominio geográfico, deciles y según fuente de **ingreso**

Cerna Maguilla, Arturo Homero (Pontificia Universidad Católica del Perú, 2023-05-08)

Si bien el crecimiento económico percibido en las dos primeras décadas del siglo XXI permitió reducir los niveles de pobreza, no logró un impacto determinante en la reducción de la **desigualdad de ingresos** de las familias. Hoy en día, los niveles de...



Estimación de un modelo "macrosocial" para la economía peruana diseñado para simular los efectos de escenarios de crecimiento y de política económica sobre la pobreza y la distribución

García Cerpo, Juan Manuel (Pontificia Universidad Católica del Perú, 2021-11-05)

El presente documento detalla la elaboración, estimación y utilización de un modelo denominado "macrosocial" para la economía peruana, que permite simular efectos a nivel microeconómico - específicamente sobre indicadores de pobreza y **desigualdad**.

LISTAR

Todo Nueva PUCP

Grados académicos y colecciones

Por fecha de publicación

Autores

Títulos

Temas

Esta colección

Por fecha de publicación

Autores

Títulos

Temas

MI CUENTA

Acceder

Registrar

DESCUBRE

Autor

Aguilar Llancor, Diego Alonso (1)

Alvarado Sazon, Anita Milagros (1)

Apaza Nina, Nancy Eliana (1)

Barrantes Gamboa, Nicolás (1)

Boyd León, Chris Marilyn (1)

... más

Tema

[https://pup.org-pe/apalcodeford5.02.01](https://pup.org.pe/apalcodeford5.02.01) (54)

Distribución del ingreso-Perú (4)

Pobreza-Perú (4)

¿Cómo se compara con el actual sistema?

Dame ejemplos de tesis sobre desigualdad de género



Sí, existen varias tesis sobre desigualdad de género en la base de datos de la Pontificia Universidad Católica del Perú (PUCP). Aquí hay algunos ejemplos:

1. "Brecha salarial de género y violencia doméstica contra la mujer en el Perú: un enfoque empírico, 2007 - 2017" por Diego Quispe Ortogorin. En esta tesis se analiza la brecha salarial de género y la violencia doméstica contra las mujeres en el Perú durante el período 2007-2017. El asesor de esta investigación es Pedro Andrés Toribio Topiltzin Francke Ballvé.
2. "Educación y Brecha Salarial de Género en el Perú" por Pamela Otilia Ríos Cahuas. Esta tesis investiga la relación entre la educación y la brecha salarial de género en el Perú. El asesor de la investigación es Luis Javier García Núñez.
3. "Informe del proyecto DIGITAGRO: Invertir en tecnología digital para aumentar el acceso al mercado de las mujeres agroempresarias en Guatemala" por Glenn Junior Kevin Ortiz Quiroz. Esta investigación se enfoca en el acceso al mercado de las mujeres agroempresarias en Guatemala y cómo la tecnología digital puede ayudar a reducir la desigualdad de género. El asesor de esta tesis es Jose Carlos Orihuela Paredes.

Ten en cuenta que los resúmenes proporcionados son fragmentos de los documentos y puede haber más detalles y hallazgos en las tesis completas. Si estás interesado en alguno de estos temas, te sugiero revisar los documentos completos para obtener más información y posibles ideas para tu propia investigación.

¿Cómo se compara con el actual sistema?

¿Qué nuevas avenidas de investigación existen sobre desigualdad de género?



Las avenidas de investigación sobre desigualdad de género pueden incluir los siguientes aspectos:

1. Brecha salarial de género a nivel regional: Según el Documento #1, existe un vacío en la literatura respecto al enfoque regional de la brecha salarial de género. Una posible investigación podría centrarse en analizar cómo varía la brecha salarial entre hombres y mujeres en diferentes regiones del Perú y qué factores podrían contribuir a estas diferencias.
2. Libertades políticas y género: Según el Documento #2, algunas investigaciones han resaltado la importancia de las libertades políticas en el desarrollo. Un enfoque interesante podría ser examinar cómo el acceso y ejercicio de estas libertades políticas difieren entre hombres y mujeres, y cómo estas diferencias pueden influir en la desigualdad de género.
3. Bienestar multidimensional y brechas de género: El Documento #3 menciona la existencia de brechas de género en términos de bienestar multidimensional. Una posible avenida de investigación sería explorar en qué dimensiones específicas del bienestar (salud, educación, vivienda, etc.) existen estas brechas de género y cuáles son los factores que las perpetúan.
4. Habilidades no cognitivas y brechas salariales de género: El Documento #4 menciona la relación entre habilidades no cognitivas y la brecha salarial de género. Un enfoque de investigación prometedor sería investigar cómo las diferencias en habilidades no cognitivas entre hombres y mujeres pueden contribuir a la persistencia de las brechas salariales de género en el Perú.

Estas son solo algunas ideas para nuevas avenidas de investigación sobre desigualdad de género, basadas en el contexto proporcionado. Recuerda que es importante revisar los documentos completos y la literatura existente para obtener más detalles y desarrollar temas de investigación con mayor profundidad.

Determinación de Documentos Relevantes

Algoritmo de búsqueda por Máxima Relevancia Marginal (MMR):

- Paso 1: Búsqueda por Similitud. Encuentra los 20 vectores más cercanos al vector que representa nuestra consulta.
- Paso 2: MMR. Escoge los 4 vectores que tengan alta relevancia con la consulta pero baja cercanía entre ellos.

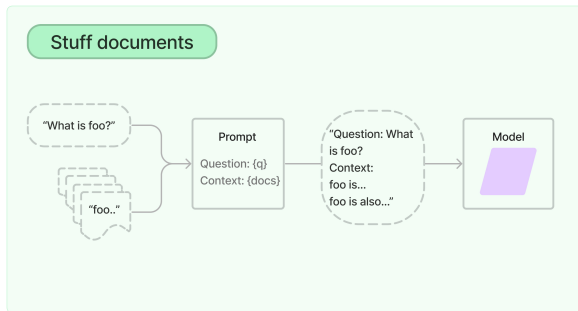
Nos encaminamos a mejorar este sistema, con mayor información por vector, más vectores en ambos casos, y optimización de los parámetros de la búsqueda.

Análisis de los documentos

El modelo de lenguaje es provisto por OpenAI. Existen varias arquitecturas posibles para integrar el modelo de lenguaje a nuestra información. Para nuestra aplicación, dos clases de arquitectura son afines:

La primera es llamada Stuff (rellenado):

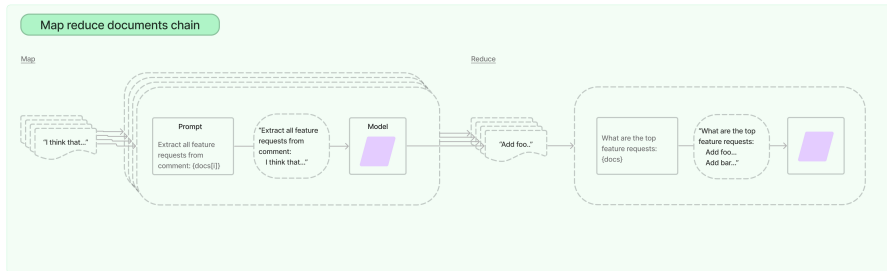
- Los fragmentos hallados son provistos de una sola vez al modelo
- Bajo costo
- Maneja pocos fragmentos de baja longitud.



Análisis de los documentos (cont.)

La segunda arquitectura es llamada Map-reduce (rellenado):

- Paso 1: Cada fragmento es analizado por el modelo independientemente para responder a la consulta
- Paso 2: Respuestas son condensadas en una sola por modelo
- Logra analizar varios fragmentos, cada uno puede ser extenso.
- Costoso



Nuevas avenidas de investigación

- Buscamos que el bot aconseje a los tesisistas sobre las posibles vías de investigación para sus tesis.
- Esto puede ser complicado puesto que requiere creatividad que escapa a las capacidades del modelo.
- Sin embargo, las tesis suelen contener discusiones sobre nuevas vías de investigación. Esta información es hallada por la búsqueda de documentos relevantes cuando la consulta lo requiere.

Principios éticos

Buscamos respetar los principios legales y éticos al desarrollar la aplicación:

- Todas las tesis son de aquellas acceso público según ha sido dispuesto por los autores.
- Nos hemos cerciorado de que el modelo tenga en cuenta siempre dar crédito a los autores cada vez que exprese una idea contenida en una tesis.

Gracias

Anexo: Procedimiento MMR

Los fragmentos D_i son hallados al resolver el iterativamente siguiente problema de maximización:

$$\max_{D_i \in R \setminus S} \left[\lambda \text{Sim}_1(D_i, Q) - (1 - \lambda) \max_{D_j \in S} \text{Sim}_2(D_i, D_j) \right]$$

Donde:

- C es el universo de fragmentos en la base
- $R \subseteq C$ es un conjunto que contiene a los elementos de C encontrados por Búsqueda por Similitud
- $S \subseteq R$ es un conjunto que contiene a los elementos previamente encontrados por MMR
- $\lambda \in [0, 1]$ es un parámetro de peso: mayores valores implican menor peso a la diversidad
- Q es la consulta vectorizada
- D_i, D_j son fragmentos vectorizados