AI Challenge

Raghavan

Arxiv Deep Dive

Help beginner researchers learn the domain and identify prominent authors

lab lab ai

Problem statement

A new researcher to a niche field, faces following problems

1. What are the jargons ? what are some key concepts i should learn? How can i understand ?
2. Who are the researchers in that niche field ? What are their papers that i can read?

Solution: arXiv deep dive

- arXiv has 2 million articles since the beginning.
- I indexed all the articles upto August 2023 using the dataset on Kaggle
- Created a vector data store using SBERT embedding.
- Used GPT 3.5 16 K (with long context) and specialized prompt to solve for the two problems

I have uploaded the vector annoy index containing all the arxiv articles (title+abstract) onto Amazon S3 for public consumption

- Annoy Index for (Arxiv articles upto 12/2022)
- Annoy Index for (Arxiv articles since 12/2022 to 08/2023)
- Arxiv data with metadata json file (upto 12/2022)
- Arxiv data with metadata json file (since 12/2022 to 08/2023)
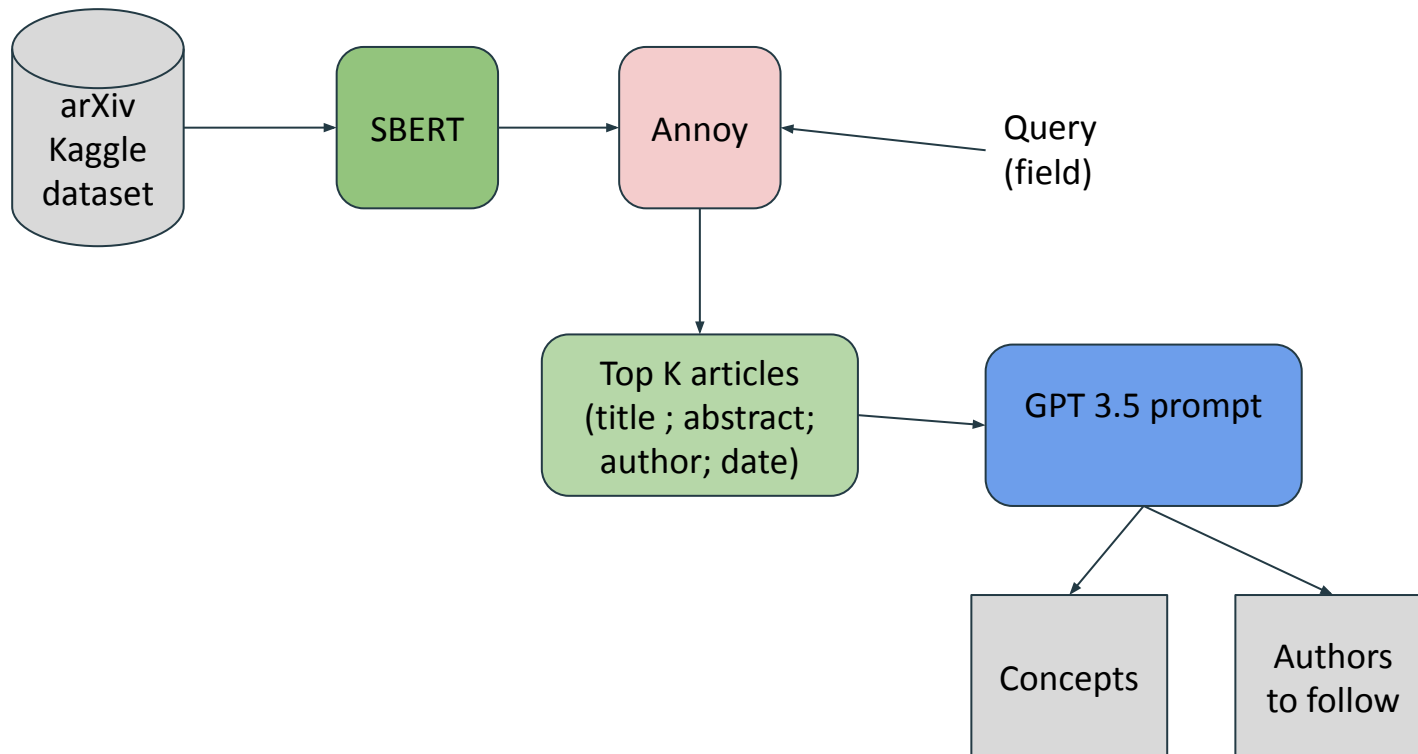
Economies of small countries

Authors you can follow on ARXIV

- **authors:**
  - **name:** Ivan O. Kitov
    - **id:**
      - 0904.0729
    - **summary:** Accurate measurement in economics according to high standards of natural sciences.
  - **name:** Austin Gerig
    - **id:**
      - 1002.0377
    - **summary:** Applying methods from natural sciences to economics and financial markets.

- **concepts:**
  - **keyword:** Gross Domestic Product (GDP)
    - **definition:** 📊 The total value of all goods and services produced in a country in a specific time period. It is used as a measure of a country's economic output.
    - **example:** Example: The GDP of a small country increased by 3% last year.
    - **analogy:** Analogy: Think of GDP as the total score of a basketball team in a game. It gives you an idea of how well the team is performing overall.
  - **keyword:** Income Inequality
    - **definition:** 💹 The unequal distribution of income among the population in a country. It measures the gap between the richest and poorest individuals or households.
    - **example:** Example: The income inequality in the country is high, with the top 1% earning a significant portion of the total income.
    - **analogy:** Analogy: Imagine a classroom where some students have a lot of candy while others have very little. This represents income inequality.
  - **keyword:** Trade Openness
    - **definition:** 🌐 The extent to which a country participates in international trade and allows the movement of goods and services across borders.
    - **example:** Example: The small country has a high level of trade openness, with a large portion of its GDP coming from exports.
    - **analogy:** Analogy: Trade openness is like opening the doors of a house to invite guests from other countries. It allows for exchange and interaction.
  - **keyword:** Foreign Direct Investment (FDI)
    - **definition:** 🏦 The investment made by a company or individual from one country into a business in another country. It involves controlling ownership and long-term presence.
    - **example:** Example: The small country attracted a lot of FDI, leading to the establishment of new factories and the creation of jobs.
    - **analogy:** Analogy: Imagine a person from one neighborhood buying a house in another neighborhood and starting a business there. This represents FDI.
  - **keyword:** Economic Growth
    - **definition:** 📈 The increase in the total value of goods and services produced in an economy over time. It is often measured as the percentage change in GDP.
    - **example:** Example: The small country experienced rapid economic growth of 5% per year, leading to improved living standards.
    - **analogy:** Analogy: Economic growth is like a tree growing taller and producing more fruits over time. It represents an increase in the overall size and productivity of the economy.
  - **keyword:** Small and Medium-sized Enterprises (SMEs)
    - **definition:** 🏭 Businesses that have a limited number of employees and a relatively small contribution to the overall economy. They are often considered the backbone of the economy.
    - **example:** Example: SMEs play a crucial role in the small country's economy, accounting for a significant portion of employment and GDP.
    - **analogy:** Analogy: SMEs are like small fish in a pond that collectively make up a large part of its ecosystem. They may be small individually, but their combined impact is significant.
  - **keyword:** Labor Force Participation
    - **definition:** 🧑 The percentage of the working-age population that is either employed or actively looking for employment.
    - **example:** Example: The small country has a high labor force participation rate, indicating a large portion of the population is engaged in the labor market.
    - **analogy:** Analogy: Labor force participation is like a game of musical chairs, where people actively participate in finding a seat (job) or preparing to find one.
  - **keyword:** Foreign Aid
    - **definition:** 🤝 Financial assistance, resources, or support provided by one country to another as a form of aid, usually to promote economic development or address specific needs.
    - **example:** Example: The small country relies heavily on foreign aid to fund its infrastructure projects and social programs.
    - **analogy:** Analogy: Foreign aid is like a helping hand given by a friend in times of need. It provides support and resources to help overcome challenges.
  - **keyword:** Tourism
    - **definition:** 🏖 The practice of traveling for pleasure or leisure to different places, often involving the visitation of tourist attractions and the spending of money on accommodations, food, and entertainment.
    - **example:** Example: Tourism plays a vital role in the small country's economy, generating revenue and creating jobs in the hospitality and service sectors.
    - **analogy:** Analogy: Tourism is like going on a vacation to explore new places and experience different cultures. It contributes to the economy of the destination.
  - **keyword:** Fiscal Policy
    - **definition:** 📜 The use of government spending and taxation to influence the economy, particularly in managing levels of aggregate demand, employment, and inflation.
    - **example:** Example: The small country implemented expansionary fiscal policy by increasing public spending to stimulate economic growth during a recession.
    - **analogy:** Analogy: Fiscal policy is like adjusting the thermostat in a room to control the temperature. It involves government actions to influence the overall economy.

Solar flares

127.0.0.1:5000/search?field=solar%20flares

Authors you can follow on ARXIV

- **authors:**
  - **name:** Brian R. Dennis
    - **id:**
      - 1109.5831
    - **summary:** Contributed to understanding particle acceleration mechanisms in astrophysical plasmas and provided an overview of solar flares in the context of observational data.
  - **name:** Hugh S. Hudson
    - **id:**
      - 1109.5831
      - 1401.6474
    - **summary:** Investigated solar flare productivity and analyzed the solar flare index for solar cycles, providing insights into the physical processes of solar flares.
  - **name:** Silvina Guidoni
    - **id:**
      - 2302.11549
    - **summary:** Contributed to the understanding of quasi-periodic pulsations in solar flares, identifying them as a key diagnostic tool for studying energy release on the Sun.
- **concepts:**
  - **keyword:** Solar Flare
    - **definition:** A solar flare is a sudden and intense release of energy in the Sun's atmosphere, often accompanied by a burst of radiation and the ejection of solar material into space. *Solar flare* is a common term used to describe this phenomenon.
    - **example:** An example of a solar flare is the X9.3-class solar flare that occurred on September 6, 2017, which was the most powerful flare in over a decade.
    - **analogy:** An analogy to understand a solar flare is like a huge explosion on the Sun, similar to a bomb detonating on Earth, releasing a large amount of energy and causing a sudden burst of light and heat.
  - **keyword:** Coronal Mass Ejection (CME)
    - **definition:** A coronal mass ejection (CME) is a massive eruption of solar material, such as plasma and magnetic fields, from the Sun's corona. *Coronal mass ejection* is a specific term used to describe this type of solar event.
    - **example:** An example of a coronal mass ejection is the Carrington Event, which occurred on September 1-2, 1859, and produced a large-scale geomagnetic storm on Earth.
    - **analogy:** An analogy to understand a coronal mass ejection is like a massive cannon shooting out a large projectile from the Sun's corona into space, carrying solar material and magnetic fields with it.
  - **keyword:** Quasi-periodic Pulsations (QPPs)
    - **definition:** Quasi-periodic pulsations (QPPs) refer to the rhythmic oscillations or variations in the emission of radiation during a solar flare. *Quasi-periodic pulsations* is a specific term used to describe these oscillatory behaviors.
    - **example:** An example of quasi-periodic pulsations is the regular fluctuations in X-ray emission observed during a solar flare, with a period of a few seconds to minutes.
    - **analogy:** An analogy to understand quasi-periodic pulsations is like a light bulb that flickers on and off in a regular pattern, with the intervals between each flicker being almost equal.
  - **keyword:** Flare Energy Release
    - **definition:** Flare energy release refers to the process by which energy stored in the Sun's magnetic field is suddenly released in the form of radiation and the acceleration of particles. *Flare energy release* is a term used to describe this release of energy.
    - **example:** An example of flare energy release is the rapid increase in X-ray and gamma-ray emissions during a solar flare, indicating the release of a large amount of energy.
    - **analogy:** An analogy to understand flare energy release is like a tightly wound spring that suddenly unwinds, releasing its stored energy and causing a rapid movement or action.
  - **keyword:** Solar Magnetic Field
    - **definition:** The solar magnetic field is a magnetic field that is generated by the motion of charged particles in the Sun's outer layers, primarily in the convection zone. *Solar magnetic field* is a term used to describe this magnetic field.
    - **example:** An example of the solar magnetic field is the complex network of magnetic loops and sunspots observed on the Sun's surface, which indicate the presence of strong magnetic fields.
    - **analogy:** An analogy to understand the solar magnetic field is like the invisible force field around a magnet, which can attract or repel other magnets and objects in its vicinity.
  - **keyword:** Geomagnetic Storm
    - **definition:** A geomagnetic storm is a disturbance in Earth's magnetic field caused by the interaction of solar particles and magnetic fields with the Earth's magnetosphere. *Geomagnetic storm* is a specific term used to describe this phenomenon.
    - **example:** An example of a geomagnetic storm is the Carrington Event in 1859, during which a strong solar flare and coronal mass ejection caused widespread disruption to telegraph systems and auroras visible at low latitudes.
    - **analogy:** An analogy to understand a geomagnetic storm is like a fierce wind blowing through a forest, causing the trees to sway and the leaves to rustle, disrupting the normal calm and creating a temporary disturbance.
  - **keyword:** Solar Cycle
    - **definition:** The solar cycle is a periodic variation in the Sun's magnetic activity, including the number of sunspots, flares, and other solar events, over a roughly 11-year cycle. *Solar cycle* is a term used to describe this regular pattern.
    - **example:** An example of the solar cycle is the transition from a period of low solar activity, known as solar minimum, to a period of high activity, known as solar maximum, and back to low activity again.
    - **analogy:** An analogy to understand the solar cycle is like the changing of the seasons, with the Sun going through periods of summer-like activity with many sunspots and flares, and periods of winter-like activity with fewer sunspots and flares.
  - **keyword:** Solar Eruption
    - **definition:** A solar eruption is a sudden release of energy and solar material, such as plasma and magnetic fields, from the Sun's surface or atmosphere. *Solar eruption* is a broad term used to describe various types of explosive events.
    - **example:** An example of a solar eruption is a solar flare, which is a localized eruption of energy and material from a region on the Sun's surface.
    - **analogy:** An analogy to understand a solar eruption is like a volcanic eruption on Earth, where molten lava and gases are violently expelled from a volcano, creating a fiery explosion and sending debris into the surrounding area.
  - **keyword:** Heliophysics
    - **definition:** Heliophysics is the study of the Sun, its effects on the solar system, and the interaction between the Sun, solar wind, and the Earth's magnetosphere. *Heliophysics* is a term used to describe this interdisciplinary field of research.
    - **example:** An example of heliophysics is the study of how solar flares and coronal mass ejections can impact Earth's magnetic field and cause geomagnetic storms.
    - **analogy:** An analogy to understand heliophysics is like studying the relationship between a star and the planets in a solar system, including how the star's activity can affect the environment and conditions on the planets.
  - **keyword:** Solar Atmosphere
    - **definition:** The solar atmosphere is the outermost layer of the Sun, consisting of several distinct regions including the photosphere, chromosphere, and corona. *Solar atmosphere* is a term used to describe these different layers of the Sun's outer atmosphere.
    - **example:** An example of the solar atmosphere is the corona, which is the outermost region of the Sun's atmosphere that is visible during a total solar eclipse as a faint, glowing halo of plasma.
    - **analogy:** An analogy to understand the solar atmosphere is like the layers of Earth's atmosphere, from the surface to the uppermost regions, each with its own characteristics and composition.
  - **keyword:** Photosphere
    - **definition:** The photosphere is the visible surface of the Sun, where most of the Sun's light and heat is emitted. *Photosphere* is a term used to describe this lowest and densest layer of the solar atmosphere.
    - **example:** An example of the photosphere is the region of the Sun's surface that appears as a bright and granulated pattern, known as granules, in images taken with telescopes.
    - **analogy:** An analogy to understand the photosphere is like the outer layer of a boiling pot of water, where bubbles form and rise to the surface, releasing heat and creating a visible pattern.
  - **keyword:** Magnetosphere
    - **definition:** The magnetosphere is the region of space around a celestial body, such as a planet or a star, where the magnetic field dominates the behavior of charged particles. *Magnetosphere* specifically refers to the Earth's magnetic field and its interaction with the solar wind.
    - **example:** An example of the magnetosphere is the region around Earth where the Earth's magnetic field deflects and traps particles from the solar wind, creating the Van Allen radiation belts.
    - **analogy:** An analogy to understand the magnetosphere is like an invisible shield that surrounds Earth, protecting it from the harmful effects of the solar wind and other space weather phenomena.
  - **keyword:** Solar Wind

Large language models

Authors you can follow on ARXIV

- **authors:**
  - **name:** Ciprian Chelba
    id:
    - 1210.8440
    - 1312.3005
    summary: Leading author in large language models with expertise in automatic speech recognition.
    reason: Ciprian Chelba has published multiple papers in the field of large language models and has expertise in automatic speech recognition, an important application of these models.
  - **name:** Adam Roberts
    id:
    - 2212.03551
    - 2212.03551
    summary: Author specializing in scaling up models and data with expertise in large language models.
    reason: Adam Roberts has published recent papers on scaling up models and data in large language models, contributing to the field's advancements.
  - **name:** Samuel R. Bowman
    id:
    - 2304.00612
    summary: Expert in large language models with focus on their surprising behaviors and limitations.
    reason: Samuel R. Bowman's work explores surprising aspects of large language models, shedding light on their capabilities and limitations.

- **concepts:**
  - **keyword:** Language Model
    definition: A Language Model is a statistical model that predicts the probability of a sequence of words occurring in a given context.
    example: Example: The language model predicts the probability of the sentence 'I love pizza' occurring in a conversation about food.
    analogy: Analogy: Just like a weather forecast predicts the probability of rain based on the current weather conditions, a language model predicts the probability of a sequence of words based on the context.
  - **keyword:** Transformer
    definition: A Transformer is a deep learning model architecture that is used in large language models to efficiently process sequential data.
    example: Example: The GPT-3 model is based on the Transformer architecture.
    analogy: Analogy: Think of a transformer like a powerful machine that can quickly analyze and understand a large amount of text, just like how a transformer in the power grid efficiently transfers electricity.
  - **keyword:** Pretraining
    definition: Pretraining is the initial phase of training a large language model where it is exposed to a large amount of general text data to learn the structure and patterns of language.
    example: Example: Before fine-tuning, GPT-3 undergoes a pretraining phase on a massive dataset of internet text.
    analogy: Analogy: Pretraining is like learning the basics of a new subject before diving into more advanced topics. Similar to how you learn the fundamentals of math before solving complex equations.
  - **keyword:** Fine-tuning
    definition: Fine-tuning is the second phase of training a large language model where it is trained on a specific task or domain by using a smaller dataset.
    example: Example: After pretraining, GPT-3 is fine-tuned on a specific task like text generation or question answering.
    analogy: Analogy: Fine-tuning is like practicing specific exercises after learning the basics. Just like how a musician fine-tunes their performance by rehearsing specific pieces of music.
  - **keyword:** Attention Mechanism
    definition: An Attention Mechanism is a component in the Transformer model that allows the model to focus on different parts of the input sequence when making predictions.
    example: Example: The attention mechanism in GPT-3 helps the model focus on relevant words when generating a response to a question.
    analogy: Analogy: Imagine you're reading a book and your attention automatically shifts to the important words or phrases. An attention mechanism in a language model works similarly, focusing on the important parts of the input sequence.
  - **keyword:** Contextual Embeddings
    definition: Contextual Embeddings are representations of words or phrases in a language model that capture their meaning based on the surrounding context.
    example: Example: In GPT-3, each word is assigned a contextual embedding that takes into account the words that come before and after it.
    analogy: Analogy: Contextual embeddings are like puzzle pieces that only make sense when they are placed in the right context. Similar to how a single word can have different meanings depending on the sentence it is used in.
  - **keyword:** Transfer Learning
    definition: Transfer Learning is a technique where knowledge learned from one task or domain is applied to another task or domain.
    example: Example: GPT-3 uses transfer learning by pretraining on a large corpus of text and then fine-tuning on specific language tasks like translation or summarization.
    analogy: Analogy: Transfer learning is like using the knowledge you gained from playing chess to improve your performance in other strategy games. The skills and strategies you learned in one game can be transferred to another.
  - **keyword:** Generative Model
    definition: A Generative Model is a type of model that can generate new samples that are similar to the training data it has learned from.
    example: Example: GPT-3 is a generative language model that can generate coherent and contextually appropriate text based on the input it receives.
    analogy: Analogy: Think of a generative model like an artist who can paint new and unique artworks based on the inspiration and techniques learned from studying different artistic styles.
  - **keyword:** Beam Search
    definition: Beam Search is a technique used in language generation tasks to find the most likely sequence of words by considering multiple candidate sequences simultaneously.
    example: Example: When generating a response, GPT-3 uses beam search to explore different word sequences and select the most probable one.
    analogy: Analogy: Beam search is like exploring different paths in a maze to find the shortest route. Just like how you consider multiple options before making a decision.
  - **keyword:** Zero-Shot Learning
    definition: Zero-Shot Learning is a technique where a language model is trained to perform tasks it has never seen before by using descriptions or prompts at inference time.
    example: Example: GPT-3 can perform zero-shot learning by generating coherent responses to prompts even on tasks it was not explicitly trained on.
    analogy: Analogy: Zero-shot learning is like solving a puzzle using only the picture on the box as a guide. You can solve the puzzle without any prior experience with it.

Thank you

Github: https://github.com/Raghavan1988/learn_a_field

2 Python files:

- Generate SBERT embeddings:
- Search for great authors and concepts

Please let me know if you have questions

1. https://www.linkedin.com/in/raghavanmit/
2. Discord: rm3844
3. Lablab,ai https://lablab.ai/u/@raghavan848