Damilola & Sunday

Developing a Chat-GPT assisted interaction to assess Normalized Difference Vegetation Index (NDVI) in R studio.

Overview

The Normalized Difference Vegetation Index (NDVI), derived from red and near-infrared spectrometric data, gauges vegetation health. Ranging from -1 to 1, it indicates barrenness (0), healthy growth (1), and non-vegetated areas (<0). NDVI can serve as a simple and reliable indicator of vegetation status, facilitating informed decision-making in various sectors. The rationale behind this is that live green plants absorb solar radiation in the photosynthetically active radiation (PAR) region for photosynthesis. They reflect near-infrared light to prevent overheating, appearing dark in PAR and bright in near-infrared. This contrast aids in distinguishing them from clouds and snow in satellite images using NDVI, calculated from NIR and red wavelengths. Accessing the NDVI dataset from NASA's earth database allows users to retrieve images for various locations efficiently. To streamline this process, automating image downloads using R saves time and ensures accuracy. Integrating ChatGPT to generate code facilitates accessing the NASA portal with user credentials (username and password) and downloading images specified by URLs stored in a text file. The research question is how to develop an automated method, potentially using R programming language and ChatGPT, to download NDVI images from NASA's earth database for specific locations or regions. For this, we concentrated on Gainesville, Florida. The primary objective is to improve efficiency and streamline the data retrieval process for research or analysis purposes. The secondary objective is to assess how right and error-free ChatGPT’s responses in making codes would be. To achieve this, a conversation with ChatGPT was started with “*I am trying to download NDVI dataset from NASA's earth data search. I chose a forested location in Florida, and I have a download link to each of the imageries. It is stored in a text file. I want to automate this download in r studio, so I do not have to download one at a time because of the inefficiency. Can you provide a working code in R studio to download these images of NDVI? I have a username and password for this NASA data portal*.” The responses and conversation are in a .txt file. After several conversations to fine-tune the code and correct errors that came up, at the **#30 conversation**, we obtained a code that worked, and the images downloaded! We tried revising the codes in R to ensure that it is automated and can be used anytime for this purpose. The different errors encountered ranged from accessing the URL to getting a correct image format to download. The case study shows how large language models (LLMs) can expedite the advancement of research workflows in environmental science when guided by critical thinking, well-defined research questions, and specific objectives. However, it will take a while to achieve the desired goal when working with LLMs for a purpose like this. Patience is essential!

References

Labib, SM; Lindley, Sarah; Huck, Jonny J. (July 2020). "Scale effects in remotely sensed greenspace metrics and how to mitigate them for environmental health exposure assessment". Computers, Environment and Urban Systems. 82: 101501.

Merow C, Serra-Diaz, JM, Enquist, BJ, & Wilson AM. 2023. AI chatbots can boost scientific coding. Nature Ecology & Evolution. https://doi.org/10.1038/s41559-023-02063-3

Owens H, Chamberlain S, Ram K, Hart T, rOpenSci. 2023. spocc: Interface to Species Occurrence Data Sources.https://cran.r-project.org/web/packages/spocc/index.html

Gates, David M. (1980) Biophysical Ecology, Springer-Verlag, New York, 611 p.

Perkel JM. 2023. Six tips for better coding with ChatGPT. Nature. https://doi.org/10.1038/d41586-023-01833-0