Luke Martin Hernandez

📍 Quezon City   ✉ luke.martin.hernandez@gmail.com   📞 09954553950

Education

Bachelor of Library and Information Science,
University of the Philippines Diliman
2019 – present
Expected graduation 2022
Specialized in Information Technology and Archiving tracks, including courses on data science and research management.

University of the Philippines Diliman
2015 – 2018
Completed 82 units towards Bachelor of Science in Computer Science
Completed relevant courses including Data Structures and Algorithms, Programming Languages, and Artificial Intelligence.

Professional Experience

Freelance Python Developer, *Calamansi Pte. Ltd.*
Mar 2022 – present
- Created a pipeline for collecting large datasets through web-scraping Twitter using Twint and Twitter API.
- Developed software that scraped Twitter data and cleaned and transformed data using Pandas for long-term storage and to assist insight teams.

Skills

Languages *(Python, Javascript, C, R, Arduino code, PHP)*,
Frameworks *(Django, Flask, Jquery, Bootstrap, Laravel)*,
Databases *(MySQL, PostgreSQL)*,
Server environments *(Digital Ocean, Alicloud)*,
Web Scraping *(Scrapy, Beautiful Soup, Selenium)*

Courses

Introduction to R, *Datacamp*

Introduction to Text Analysis in R, *Datacamp*

Projects

PDF2MARC
2022
An expert system that leverages machine-learning for Natural Language Processing (NLP) to automatically extract descriptive cataloging information from PDFs. It is trained against a large dataset scraped from Wikipedia articles for named entity recognition as well as user-customizable rulesets.
This was created as part of my thesis for my degree in Bachelor of Library and Information Science.
Languages: Python, HTML, Javascript
Frameworks: Flask, Bootstrap 5, Jquery
Libraries: Pandas, SpaCy, NLTK, Regex
https://github.com/lhernandez0/PDF2MARC

Twitter Scraper
2022
A Twitter scraper utilizing Twint and the Twitter API to collect large datasets of tweets and users related to inputted keywords. Cleans and transforms datasets according to the user or client needs, and is exportable in a variety of formats including CSV and SQL for analysis and insight generation.
Languages: Python
Libraries: Pandas, Twint
https://github.com/lhernandez0/Twitter-Scraper

KucoinBot ↗
2021
An automatic portfolio-rebalancing bot leveraging algorithmic trading techniques. Interfaces with the Kucoin API for automated trading, and Telegram API for user interface and interaction. Visualizes users' historical portfolio value and gains to improve understanding and maximize profit.
Languages: Python
Libraries: Pandas, matplotlib
https://github.com/lhernandez0/KucoinBot

Introduction to the Tidyverse, *Datacamp*

Network Analysis in R, *Datacamp*

Visualizing Geospatial Data in R, *Datacamp*

Interactive Data Visualization with plotly in R, *Datacamp*

Research Data Management and Sharing, *Coursera*

References

Percival Tolentino, *Project Manager*, Calamansi Pte. Ltd.
pjatolentino@gmail.com, 09178975685