

Chi-Yun Chen

☎ +1 (202) 868 0865 | ✉ chiyun.chen001@gmail.com | 📱 [cychen116](#) | 🌐 [chi-yun-chen](#)

EDUCATION

The George Washington University

MASTER OF SCIENCE IN DATA ANALYTICS

Washington DC, US

Expected Dec 2022

- Courses: Machine Learning, Statistics, Big Data, Database Management, Algorithms, Marketing Analysis, etc.
- Honors: GPA 3.8, Fellowship & Best ML/AI approach prize in 2021 GW Hackathon.

National Chung Hsing University

BACHELOR OF SCIENCE IN FOOD SCIENCE AND BIOTECHNOLOGY

Taichung, Taiwan

Jan 2016

EXPERIENCE

The George Washington University

IT ADMINISTRATION DATA PROJECT ANALYST

Washington DC, US

Feb 2022 – Present

- Analyzed and evaluated the IT data by Excel Pivot Table, meanwhile generating a performance dashboard with insight.
- Collected, maintained, and wrangled computer status data verification and validation by Python among 6,000 staff and 40 departments.

iChemLabs LLC

DATA INTERN

Remote, US

Jul 2020 – Sep 2021

- Evaluated thousands chemical compounds data and wrote testing cases in the Java environment.

National Yang Ming University

PHARMACEUTICAL RESEARCH ASSISTANT

Taipei, Taiwan

Jul 2018 – Aug 2019

- Implemented data analysis on diverse biotechnology research of nanoparticles in cancer therapy by Excel and R.
- Researched scientific literatures of pharmaceuticals to pull out required data.
- Leveraged experimental data to develop research methods in custom-made quantitative frameworks.

Innovation Open House

DATA ANALYST AND PRODUCT MANAGER

Taipei, Taiwan

Jan 2016 – Jul 2017

- Inspected website performance data by GA and executed SEO resulting in 74% sessions increasing.
- Developed MySQL database from scratch and systematized data storage by building NAS for a startup company.
- Investigated customers' needs and implemented A/B testing for video products, resulting in a 200% average view duration improvement.
- Analyzed products data by MySQL, then created an automatic performance evaluation system resulting in a 20% turnaround time reduction.
- Utilized statistical methods and Excel pivot table to execute marketing campaign and increased engagements of the FB Fan Page by 30%.
- Collaborated and led a cross-functional team 3 times for a live-stream event that owns tens of thousands of audiences.
- Designed and maintained KPI dashboards among 20 interns and led new employee orientation.

SKILLS

Programming Languages: Python (*NumPy, pandas, SciPy, TensorFlow, PyTorch, scikit-learn, etc.*), R (*Dplyr, Knitr, RMarkdown, Shiny, etc.*), SQL.

Data Science & Miscellaneous Technologies:

- Data Science Pipeline: ETL Pipeline, Cleaning, Wrangling, Exploratory Data Analysis, Modeling, Interpretation.
- Data Visualization: Python (*matplotlib, seaborn*), R (*ggplot2*), Tableau, QGIS.
- Statistics Analysis: Regression Analysis, ANOVA, PCA, Time Series, A/B Testing, MATLAB, GA, Excel.
- Others: Machine Learning, Deep Learning, Natural Language Processing, Text Processing, Topic Modeling.

Big Data & Database Management: MySQL, MongoDB, Neo4j, MLaaS (*AWS, Google Cloud*).

PROJECTS

Sign Language Translator by Machine Learning Algorithm

bit.ly/SL_trans

- Regulated motion capture and computer vision recognition module by OpenCV and Mediapipe libraries.
- Built Deep Learning models by PyTorch, such as LSTM, RNN, and CNN, and studied sign language structure.

Analysis of COVID-19 Topic on Facebook/Twitter in LDA Evaluation Metrics by PCA

bit.ly/LDA_dan_21

- Extracted social media data by CrowdTangle and performed an analysis in Natural Language Processing by Python.
- Adopted MALLET toolkit to conduct topic modeling and Latent Dirichlet Allocation (LDA).

Exploratory Data Analysis Repository Using Statistical and Machine Learning Methods

bit.ly/EDA_github

- Utilized data science pipeline and ETL process in various dataset from Kaggle.
- Adopted time series analysis, regression analysis, classification analysis, prediction modeling and model optimization.

Exploratory Data Analysis and Recommendation System of Yelp Dataset Management

bit.ly/yelp_git

- Restructured Yelp dataset into SQL server by AWS and Python SQL libraries.
- Constructed a Yelp Recommendation System by Machine Learning and Deep Learning models.

Machine Learning Approaches in Drug Discovery

bit.ly/drugdiscover_git

- Extracted target protein data from the ChEMBL dataset and pre-processed and transformed the data.
- Built a drug discovery model by statistical and Machine Learning models, then created a web application to deploy the final model.