Wayland La

linkedin.com/in/waylandla|github.com/Dnalyaw | dnalyaw.github.io | (415) 350-3607 | waylandla@berkeley.edu

EDUCATION

University of California, Berkeley   **Expected Graduation: May 2026**
B.A. Data Science, B.A. Applied Mathematics   **GPA: 3.9 / 4.0**
Relevant Courses: Structure of Computer Programs, Foundations of Data Science, Multivariable Calculus, Linear Algebra
Organizations: Data Science Foundations Scholar, Data Science Discovery Research Program, Data Science Society

SKILLS & INTERESTS

- **Technical Skills:** Python, Java, SQL, Microsoft Excel, Microsoft Power-Bi, Pandas, NumPy, RegEx, Matplotlib, Seaborn, Scikit-learn, PyTorch, TensorFlow, Keras, NLTK, HTML, CSS, Javascript, Jupyter, Google Colab, Anaconda, Tableau
- **Soft Skills:** Leadership, Problem-Solving, Strategic Planning, Strong Communication, Critical Thinking, Team Player

WORK EXPERIENCE

Data Science Intern (UC Berkeley CDSS Data Discovery Research Extension)   Sep 2024 – May 2025
Oakland Natives Give Back   *Oakland, CA*
- Created ML models (**Random Forest Classifier, XGBoost Classifier**) as well as implemented simple Deep Learning Neural Networks (**Tensorflow**) to predict whether Oakland public school students were at risk of being chronically absent
- Cleaned empty data and used Feature engineering on messy dataset through Pandas and used EDA such as pie histograms, heatmaps and scatterplots to find correlational features for students who are likely to become absent in the future
- Recommended & Implemented Programs targeting chronically absent students to improve the rate of attendance

UC Berkeley, College of Computing, Data Science, and Society   Aug 2024 – May 2025
DATA C8 UCS1 - Tutor   *Berkeley, CA*
- Supporting 15+ student tutoring sections with assignments in an foundational data science course of 1100+ students
- Conducting thorough content reviews for assignments, labs, and projects with **statistics/data analysis principles**
- Constructing exam-level problems for the course's midterm and final exams to assess student understanding, and facilitating exam proctoring by monitoring activity and enforcing academic integrity for a secure testing environment

Data Science Academic Development Mentor   Sep 2024 – May 2025
Data Science Society   *Berkeley, CA*
- Teaching fundamental ML topics (**Regression, Classification, Supervised/Unsupervised Learning**) and EDA topics (Pandas, Matplotlib, Seaborn) and in an Introduction to Real World Data Science course open to 70+ students
- Mentored students with creating full scale real world data science/ML projects centered around the Data Science Lifecycle

PROJECTS

CNN-Chatbot (RAG Model) | *LLMs (GPT 3.5 Turbo), Langchain, HuggingFace (Embeddings, Bias Classifier)*   Aug – Dec 2024
- A summarization tool catered to CNN articles that can contextualize its responses to queries on a previously stated topic
- Uses cosine similarity embeddings and web scraping to find best fit CNN articles to answer based on the topic given
- Chunks and tokenizes data into OpenAI GPT 3.5 Turbo to train the chatbot to answer questions based on new data given
- Chatbot gives descriptive answer to the prompt, and sites which CNN articles the Chatbot processed to receive answer

Student Absenteeism Prediction Model | *Tensorflow/Keras Neural Networks, Machine Learning (XGBoost)*   Aug – Dec 2024
- Built an End-to-End Data Science Project as an ONGB intern studying important features that correlate to student absenteeism in Oakland Public schools in order for the non-profit to create programs attacking students at risk
- Using artificial neural networks, created >80% accurate diagnosis for Parkinson's Disease, comparing it to XGBoost ML
- Includes visualizations (histograms/heatmaps) showing significant differences between absent/non absent students

Noise Pollution in the US Analysis | *Machine Learning w/ Sklearn, Hypothesis Testing, EDA*   Jan– Apr 2024
- Studied the correlations between certain impactful features and the amount of Excess Noise Pollution recorded in a state
- Worked on creating detailed choropleth maps with other Berkeley Undergraduates to display difference based on state
- Hypothesis tests displaying correlations between excess noise and population density, shown by histograms & scatter plots
- Created Machine Learning models to predict a future city's approximate amount of excess noise based on data trained on