Abdoulfatah Abdillahi

[GitHub](#) | [Linkedin](#) | abdoulsfsu@gmail.com | Bay Area, CA| +1 317-828-1996

EDUCATION

Master of Science (MS), Data Science and Artificial Intelligence　　　　　　　　*Expected Dec 2025*
San Francisco State University — San Francisco, CA

Bachelor of Science (BS), Computer Science　　　　　　　　　　　　　　　　　*2021–2023*
San Francisco State University — San Francisco, CA

Northern Essex Community College — Haverhill, MA　　　　　　　　　　　　　*2018–2020*

INDUSTRY EXPERIENCE

Bioinformatics Intern, Genentech, South San Francisco, CA　　　　　　**June 2023 – September 2023**
Project Intern: ***Unveiling Treatment Effects from Digital Health Data through Pharmacology-Informed Neural-SDE***

- Gained hands-on experience with neural stochastic differential equation (SDE) model to analyze stochastic data from digital health technologies (DHT) during my internship.
- Demonstrated initiative by actively seeking clarification through regular discussions, improving my ability to ask insightful and targeted questions.
- The code was written in Tensorflow, I had the opportunity to switch to a PyTorch to enhance the model training efficiency. Additionally, this project also provided an opportunity to strengthen my coding skills and gain hands-on experience in bioinformatics.

RESEARCH EXPERIENCE

Graduate Student Researcher, San Francisco State University, CA　　　　　　**Sep 2023 – Present**
 Advisors: Dr. Pleuni Pennings & Dr. Sara El Alaoui
Research Project: **Machine Learning-Driven Pan-genome Pipeline for Drug Resistance Prediction in *E. coli***

- **Preprocessed genomic data** using bioinformatics tools (e.g., Fasterq-Dump, Prokka, Roary, etc.) to construct gene presence/absence matrices and integrate SNP data.

- **Developed a machine-learning pipeline** integrating gene presence/absence and SNP features within a pan-genome framework to predict antibiotic resistance in *Escherichia coli*.

- **Led a sub-team** focused on designing, training, and evaluating a Variational Autoencoder (VAE) model, which outperformed convolutional neural networks (CNNs) in predictive accuracy.

- **Investigated resistance to Amoxicillin (AMC),** demonstrating the VAE's superior performance in capturing latent genomic patterns relevant to resistance.

- **Overcame gaps in biological background** by actively collaborating with microbiology lab members to deepen domain understanding.

RESEARCH PRESENTATIONS

- Abdillahi, A., Kebebew, E., Baldonado, G.C., Barajas, J.F., Torres, M., Yoon, I., & Pennings, P. (2024, November 8).
 Machine Learning-Driven Pan-genome Pipeline for Predicting *E. coli* Drug Resistance.
 XXIII Bay Area Population Genetics Conference (BAPG), University of California, Berkeley, CA.

- Abdillahi, A., Kebebew, E., Barajas, J.F., Torres, M., & Pennings, P. (2024, October 31 – November 2).
 Machine Learning-Driven Pan-genome Pipeline for Predicting *E. coli* Drug Resistance.
 SACNAS 2024 NDiSTEM Conference, Phoenix, AZ.

- Abdillahi, A., Kebebew, E., Baldonado, G.C., Barajas, J.F., Torres, M., Yoon, I., & Pennings, P. (2024, March 6–10). **Predicting *E. coli* Drug Resistance through Different Deep Learning-Based Approaches Using a Comprehensive Pan-genome Assembly**. *Allied Genetics Conference (TAGC24)*, Washington, DC.

TEACHING EXPERIENCE & GRADER

Braven (Remote) — *Grader*   **Sep 2024 – May 2025**
- Graded assignments and projects for a cohort of 24 students using AU 220 rubrics.
- Provided detailed feedback to enhance students' career and leadership development.
- Held weekly virtual office hours to support learning, address concerns, and foster academic success.
- Reviewed instructional materials including slide decks, projects, and rubrics to ensure alignment with learning outcomes.

Mentor / Teaching Assistant
San Francisco State University
- **CSC 408: Data Science for Personalized Medicine (Python)**   *Sep 2024 – Dec 2024*
- **CSC 219: Data Structures for Data Science Application Development (Python)**   *Jan 2024 – May 2024*
- **CSC 101: Introduction to Computing and Programming (Java)**   *Jan 2024 – May 2024*

Research Peer Mentor — PINC Summer Program, SF State   *Summer 2024*
Advisor: Dr. Pleuni Pennings
- **Project Title**: *The Correlation Between Quinolone Use and ST131 E. coli Spread in European Countries*
- Mentored five students in R programming and guided them through a full research pipeline using data from the European CDC and Enterobase.
- Investigated the relationship between quinolone consumption and the prevalence of ST131 across ten European countries.
- Findings revealed no correlation, suggesting a need for multifactorial approaches to address antibiotic resistance.

HONORS & AWARDS

NIH – UCSF/SFSU Bridges to the Doctorate Fellowship, SFSU   *2023–2025*
Two-year research training fellowship awarded to MS students from underrepresented groups in science to support transition into and completion of biomedical research-focused doctoral programs.

BMC Scholarship, SFSU   *September 2024*
Awarded $1,750 by the Computer Science Department for academic excellence and commitment to research.

National Diversity in STEM (NDiSTEM) Travel Scholarship, SACNAS Conference, Phoenix, AZ   *October 2024*
Received full travel support including roundtrip flight, lodging, and conference registration to attend the 2024 NDiSTEM Conference.

Award of Recognition, SFSU PINC Program   *May 2024*
Recognized for exemplary performance and service to students and faculty, contributing significantly to the success of the PINC community.

Certificate of Appreciation, SFSU PINC Summer Research Program   *July 2024*
Acknowledged for outstanding work as a peer mentor, demonstrating excellence in leadership and student guidance.

CSU-NSF Louis Stokes Alliance for Minority Participation (CSU-LSAMP), SFSU   *2022–2023*
Selected and awarded $500 to support underrepresented students in STEM fields.

Genentech – PINC Fellowship, SFSU   *2022–2023*
Awarded $4,500 to 9 students participating in the PINC Program for excellence and potential in computational biology and AI.

VIR STEM Mentorship Program Participant, SFSU   *December 2022*
Selected for the 2023 cycle, recognizing strong academic performance and dedication to professional growth in STEM.

Certificate

Data Science and Machine Learning for Biotechnology Certificate   *2022–2023*
San Francisco State University (in collaboration with Genentech)
- Completed a 12-unit, two-semester training program designed in partnership with Genentech to address technical needs in the biotech and pharmaceutical industries.
- Gained hands-on experience in high-demand data science and machine learning skills tailored for solving complex problems in biotechnology.
- Trained in cutting-edge tools and diverse computational approaches relevant to real-world biotech applications.

LEADERSHIP & OUTREACH

Muslim Student Association (MSA), SFSU
Da'wah Coordinator | September 2024 – Dec 2024
- Organized and facilitated weekly Friday prayers on campus in collaboration with local Imams, creating a welcoming and inclusive space for Muslim students.
- Fostered a strong sense of community and belonging by supporting students in connecting with their faith and peers.

Monroe Elementary School, San Francisco, CA
Mentor | March 2024 – May 2024
- Collaborated with Dr. Pleuni Pennings and lab members to introduce basic coding concepts to 4th grade students through weekly classroom visits.
- Designed and led engaging projects like interactive name poem games to spark early interest in technology.

Professional & Community Memberships
- Member, **National Society for the Advancement of Chicanos and Native Americans in Science (SACNAS)**
- Member, **Muslim Student Association (MSA)**
- Member, **CSU-NSF Louis Stokes Alliance for Minority Participation (CSU-LSAMP)**
- Member, **National Society of Black Engineers (NSBE @ SFSU)**

TECHNICAL SKILLS

Programming: Python, Java, JavaScript, HTML, CSS, Linux (proficient), C and C++(Intermediate)
Framework & Library: PyTorch, Keras, Tensorflow, NLTK, React, Node.js, Express, Flask, Dockerfiles, MatLab, LaTeX
GitLab/Github: branches, code reviews, pull requests
Experience working with high-performance computer clusters (Slurm)
Languages: French, Somali, and Arabic

ACADEMIC PROJECTS

Energy Consumption Modeling – Regression & Clustering Project
MATH 448: Statistical Learning | San Francisco State University   *Feb 2025 – May 2025*
- Investigating the Appliance Energy Prediction dataset (UCI ML Repository) to explore relationships between environmental variables and household energy consumption.
- Applying regression techniques including linear models, decision trees, and neural networks to forecast energy usage based on temperature, humidity, and weather data.
- Using clustering algorithms to uncover patterns in consumption behavior for insights into sustainable energy management.

Creator & Maintainer – `database-mongodb` Python Package | [PyPI](#) • [GitHub](#)   *Apr 2025 – May 2025*
- Developed a lightweight Python package to automate MongoDB connections, database creation, and bulk data insertion from CSV/Excel files.
- Packaged and deployed the module to PyPI (`v0.0.11`), configured with `tox`, `setuptools`, `requirements.txt`, and `pyproject.toml`.
- Designed a modular structure with unit tests and GitHub Actions for CI/CD, ensuring maintainability and ease of installation via `pip`.

Full-Stack Developer – Educlouds App   *Jan 2024 - May 2024*

AI Entrepreneurship Course Project GitHub Repo

- **Frontend Development**: Designed and implemented a responsive user interface using modern web technologies, enabling students to easily access and request computational resources.

- **Backend Engineering**: Developed RESTful APIs and backend logic using Flask, Python, and SQLAlchemy to manage data flow and facilitate secure transactions.
- **Project Showcase**: Successfully presented the completed Educlouds platform during the end-of-semester CSC 890 showcase, demonstrating full-stack functionality and real-world applicability.

Data Visualization Project – Shopping Trends and Customer Behavior   *Jan 2024 – May 2024*

CSC 805: Data Visualization Shopping Trends

- Built an interactive web dashboard using **Dash**, **Plotly**, and **Python** to visualize and explore customer behavior from a Kaggle dataset (~3,900 records).

- Engineered modular visualizations including a **product insights bar chart**, **seasonal trends graph**, **U.S. heatmap of popular items**, and **demographic filters** (age, gender) for purchase behavior.
- Delivered a fully functional data exploration tool that enables users to dynamically filter by age, gender, and product type, offering real-time insights into customer trends and purchase patterns.

Fitness - FitBuddyPal   *Sep 2023 - Dec 2023*

Software Engineer Course Project Fitness App

- Led the integration of the ChatGPT API to automated and personalize workout and nutrition plans within the FitBuddyPal, enhancing user engagement and satisfaction
- Led Agile Practices: Acted as Scrum Master, organizing bi-weekly meetings to review team progress, address bottlonecks, and ensure on-time delivery of project milestones.

REFERENCES

Dr. Sara El Alaoui

Assistant Professor, Department of Computer Science, San Francisco State University

Thesis Supervisor

elalaoui@sfsu.edu

Dr. Pleuni Pennings

Professor, Department of Biology, San Francisco State University

Co-Thesis Supervisor

pennings@sfsu.edu

Dr. Ilmi Yoon

Teaching Professor and Director of Computing Programs, Khoury College of Computer Sciences, Northeastern University

i.yoon@northeastern.edu