Abel D Metek

Seattle, WA | +1 2064128760 | abelmetek19@gmail.com | Linkedin | Github | Portfolio

Highly Trained Data & ML Scientist | Python, SQL, Cloud Technologies | Model Creation & Deployment

SUMMARY

- Hardworking and Diligent Data/ML Scientist with demonstrated expertise in developing and deploying machine learning & NLP models. Proficient in Python, SQL, and cloud technologies, with a focus on delivering data-driven solutions. Experienced in optimizing models for performance and extracting actionable insights from complex datasets. Adept at collaborating in dynamic environments to drive impactful results. Committed to continuous learning and applying cutting-edge techniques to solve business challenges.

SKILLS & COMPETENCIES

- **Programming Languages**: Python, SQL, C++, Java experience
- **Machine Learning/Deep Learning**: TensorFlow, PyTorch, Scikit-learn, Keras, CNN, RNN, ANN, Regression, Classification, Time Series Analysis, Statistical Modeling, GPU Optimization
- **NLP:** Nltk**,** regex, Tfidf
- **Data Analysis/Visualization**: Pandas, NumPy, Seaborn, Tableau
- **Cloud/Big Data**: AWS/Azure/GCP (Basic familiarity), Docker, Hadoop (Introductory Experience)
- **Web Development/Tools**: FastAPI, Flask, Jupyter Notebook, Google Colab
- **Version Control**: Git, GitHub, GitLab
- **Emerging Technologies**: Gen AI
- **Math**: Probability, Statistics, Calculus, Linear Algebra
- **Soft Skills**: Analytical, Problem-Solving, Collaborative, Communication, Self-Directed

EDUCATION

University of Washington **Seattle, WA**
Bachelor of Science (Computer Science) September 2021 - December 2023

- Experienced working in programming languages like Java, Python, C/C++, with a strong computer science mindset and ability to learn new languages quickly.
- Experienced in data structures & algorithms, object-oriented & functional programming, software engineering principles, and algorithmic problem-solving (greedy algorithms, dynamic programming).
- Strong foundation in mathematics (Linear Algebra, Calculus, Probability, Statistics) applied to machine learning, deep learning, and explainable AI.
- Hands-on experience in database management, SQL, data warehousing, and data visualization (Tableau).
- Enhanced teamwork and communication skills through collaborative software and data-driven projects.
- Relevant Coursework: CSE 331, CSE 332, CSE 341, CSE 333, CSE 421, CSE 427, CSE 442

WORK EXPERIENCE

Oeson **Remote**
Data Science Intern November 2024 - Feb 2025

- Gained hands-on experience in Python, statistics, hypothesis testing, machine learning, deep learning, and Tableau through real-world data projects.
- Led a capstone project on time series analysis, leveraging the Prophet model to predict future monthly average temperature anomalies based on historical dataset from (1880–2024).

- Developed a complete end-to-end pipeline, from data gathering and preprocessing to model training, deployment, and continuous maintenance on Streamlit Cloud.
- Designed Tableau dashboards to provide stakeholders with actionable data insights. Github_Link, Demo

PROJECTS

POC Relevant TeamWork Projects from CourseWork CSE 442(Data Visualization) **Dec 2023**

- Developed interactive visualizations for counting and base conversions using HTML, CSS, JavaScript, and React, improving user comprehension through dynamic UI components.
- Contributed to front-end development, focusing on HTML/CSS styling and base conversion concepts, while collaborating on React and JavaScript for interactivity.
- Worked in a team of four using agile methodologies, coordinating via Discord, and delivered a proof-of-concept demo video showcasing the project (GitHub: [link], Demo: [link]).

POC End to End Calorie Burnt Regression Project **Jan 2025**

- Developed a Random Forest regression model with $R^2 = 0.92$, leveraging Scikit-learn to predict calorie expenditure.
- Preprocessed and analyzed 10K+ data points using Pandas and NumPy, ensuring data quality and feature engineering.
- Built and deployed a Streamlit application for real-time calorie predictions, enhancing user accessibility. (GitHub: [link], Demo: [link]).

POC End to End Customer Churn Classification Project **Feb 2025**

- Built a Random Forest Classifier to predict customer churn with an F1-score of 0.89, optimizing performance through cross-validation.
- Engineered features and preprocessed telecom data to enhance model accuracy and interpretability.
- Deployed a Streamlit application to provide real-time churn insights, enabling data-driven decision-making. (GitHub: [link], Demo: [link]).

POC End to End Lung Cancer CNN with Transfer Learning Project(in progress) **Mar 2025**

- Fine-tuning VGG19 using TensorFlow/Keras on 5,000+ lung CT scan images for multi-class lung cancer classification.
- Aiming to achieve 90%+ accuracy, leveraging transfer learning and advanced hyperparameter tuning.

POC Simple End to End NLP Project Resume Screening Project **Mar 2025**

- Developed a resume classification system using TF-IDF vectorization and regular expressions for text preprocessing, along with label encoding for structured output representation.
- Implemented OneVsRestClassifier on KNN, achieving an impressive 98% accuracy on the test dataset, demonstrating strong model performance for multi-class resume categorization.
- Built a fully functional Streamlit web app for real-time resume screening, enabling seamless user interaction, with a GitHub demo showcasing end-to-end deployment. (GitHub: [link], Demo: [link]).