Brooklyn, NY

Efe Sahin

efe-sahin.com efesasa@gmail.com +1 814-826-8562 linkedin.com/in/efesas github.com/Efesasa0

Current Status

Junior Programmer at Columbia Data Analytics. I am on OPT right now and currently exploring ML, LLMs, Generative AI, Computer Vision, and Brain-Computer Interfaces.

Technical Skills

Python, Apache Spark, AWS, DataBricks, Pytorch. Scikit-learn, TensorFlow, Pandas, Seaborn Numpy Matplotlib, Linux Ubuntu, Shell, Bash, SQL, MATLAB, C, Java, Swing, Apache Derby, JavaScript, DrRacket, Verilog HDL, Microsoft 365, Latex, Overleaf, Google Colab, Kaggle, Git, Docker

Work Experience

Columbia Data Analytics, Manhattan, NY, 10013 **08.2023 - Present**

As a Junior Programmer, I have been working with Databricks Apache Spark on the AWS platform. My responsibilities include utilizing PySpark and SQL to source, cleanse, and query data from several large datasets about US health insurance providers. I worked with pharmaceutical companies like Pfizer.

Projects Achievements

Hdmr-opt app2scale

Developed a wrapper function needed to optimize hyper-parameters of XGB in the context of forecasting e-commerce transaction load.

RealTime Speech Censorship

As a group, we designed a multi-threaded module to "bleep" out banned words inside a real-time feed audio stream. I also documented a model card for the OpenAI whisper model we used. Achieved output streams with reasonable delay.

Instrumented Mouthguard Design

In the Capstone group project, I searched for ways to record and transmit kinetic data inside a mouth guard. With Dr Reuben Kraft's lead from the Biomedical Engineering department at Penn State, the project got into 2nd place in the K12 awards.

Maze solver via auto-encoder

I generated a custom maze dataset and developed an image-based maze solver which utilized an auto-encoder network. Outputs drew silhouettes of the paths in red with about 90% accuracy.

32-bit pipelined CPU design

I implemented CPU design which fed on a forwarded stream of inputs and computed Addition, Subtraction, OR, AND, and XOR commands. The design was optimized for recall and speed.

Fs3 In-memory filesystem

I developed a double-linked-list based in-memory file system with cache, read, write seek, and other operations in C programming language using tools like gdb on Ubuntu Machine.

Room scheduler App

I developed a standalone Java application using Swing library for the interface design and Apache Derby databases for keeping track of scheduling room tasks.

Education

Penn State University, University Park, PA **2020 - 2023**

Completed Bachelor of Computer Science,
B.S. Engineering in 3 years with a 3.40 GPA.

International Baccalaureate, Istanbul, Turkey **2016 - 2020**

Completed IB Diploma program when attending MEF International School and took advanced courses e.g:: HL ITGS, HL Math

Certs. & Schools

Fundamental Neuroscience for Neuroimaging **01.2024**

Completed course from Johns Hopkins on structural and functional brain scanning technologies most notably MRI, fMRI, Diffusion Tensor, and Magnetic Spectroscopy Imaging by Coursera.

Data Scientist with Python **12.2023**

Completed the track of 90 hours made available by Data Camp.

Harvard University, Cambridge, MA **06.2019 - 08.2019**

Completed residential extension school program and took two undergraduate courses simultaneously:: (A)Programming with Python and (B)Computer Science with Java

Kaplan International School, Manhattan, NY, 10118 **08.2018**

Completed TOEFL and Academic English Intensive course for one month with 100% attendance.

Purdue University, West Lafayette, IN **07.2017**

Completed GERI residential summer camp and completed 4 courses, e.g:: Programming & Computational Thinking, Discovering Scientific Laws & Concepts Through Experimentation

Learning Achievements

CMPSC 497 Deep Learning for Computer Vision: Studied complex neural networks, modern ConvNets, Transformers, object detection, 3D vision, AI interpretability, and zero-shot/few-shot learning.

MATH 452 Deep Learning Algorithms and Analysis: Focused on heavy theory and light practical aspects of deep learning, integrating mathematical foundations and numerical methods. Studied practical applications using popular databases and traditional numerical methods to understand deep neural network functions.

CMPSC 473 Operating System Design and Construction: Implemented functions from scratch like malloc, free, realloc, and multithreaded programming for reading files with the C pthread library. I learned about essential concepts such as Mutex locks, semaphores, and producer/consumer queues.

CMPSC 465 Data Structures and Algorithms: Learned about fundamental computer science concepts such as data structures, algorithm analysis, recursion, trees, sets, graphs, and sorting.

CMPSC 464 Introduction to Theory of Computation: Learned heavy theoretical concepts of computation, including abstract models of computational devices, formal languages, and computability. I learned about finite automata, regular and context-free languages, Turing machines, undecidability, and time/space complexity, including P versus NP.

CMPSC 461 Programming Language and Concepts: Implemented a recursive descent parser for a simplified HTML language. Additionally, I learned and applied functional programming throughout the course by completing various assignments using DrRacket.

Language

Turkish (Native), English (Advanced),
Latin (Beginner), Spanish (Beginner)

Hobbies

Reading non-fiction books, Watching science-fiction movies, Basketball, Swimming