# EDUCATION

Almas Fathimah

667-334-1464 | [almasfathima4@gmail.com](mailto:almasfathima4@gmail.com) <https://www.linkedin.com/in/almas-fathimah-90/>

<https://github.com/AlmasFathimah124/>

**University of Maryland, Baltimore County,** Baltimore, MD CGPA: 3.7 Master of Professional Studies, Data Science *Expected, Dec 2024*

**Nimra College of Engineering and Technology,** Ibrahimpatnam, India CGPA: 7.1 Bachelor of Engineering and Technology, Electronics and Communication Engineering *Graduated, May 2012*

# SKILLS

**Programming Languages :** Python, SQL, C,C++, HTML/CSS, PL/SQL, Java

**Databases :** MySQL, Oracle, MongoDB, SQL Server

**Tools :** IBM Rational DOORS,ALM

**Analytics Tools :** SSRS, SSIS, Microsoft Excel, Tableau, SAS, Power BI, SSAS

**Big Data Technologies :** Hadoop, Hive, Spark, Kafka, YARN

**AWS Cloud Tools :** EC2, EMR, S3

**Python Packages :** Pandas, PySpark, NumPy, Matplotlib, SciPy, OpenCV, CNN models, Dash, sklearn, keras, Rest API’s

**IDE’s :** PyCharm, Jupyter Notebook, Eclipse, Visual Studio

**OS/Platforms :** Windows, Mac OS, Linux

# PROFESSIONAL WORK EXPERIENCE

**Senior Systems Engineer,** *Amadili Labs Private Limited, India* *June 2020 – Aug 2022*

* Experienced working with databases such as Oracle, SQL Server and, MySQL, PostgreSQL, writing stored procedures, functions, joins, and triggers for different Data Models.
* Transformed batch data from several tables containing thousands of records from PostgreSQL, Teradata, and CSV file datasets into data frames using PySpark.
* Instrumentally worked with peers and stakeholders to organize, access, monitor, and control the evaluated data from a variety of sources related to customer feedback, ratings, and review data using SQL.
* Collected new data, connected data types, and used existing data creatively to formulate solutions using Python and SQL. Queried the data and provided reports to stakeholders.
* Efficiently worked on Python automation scripts to streamline data processing jobs running on CA Autosys with a good understanding of Relational Database Design, Data Warehouse/OLAP concepts, and methodologies. Implemented Spark with Python and Spark SQL for faster testing and processing of data.

**Embedded Software Developer,** *UST GLOBAL TECHNOLOGIES , Bangalore, India April 2017 – Jan 2018*

* Involved in testing and providing maintenance and documentation at each stage.
* Involved in generating auto-test cases using the Cantata tool.
* Developed and implemented test cases according to the code.
* Providing boundary value checks for variables.
* Worked on LabVIEW, which provides a real-time environment.
* Excelled in finding, documenting, and reporting bugs on time.
* Mentor a junior engineer as part of his/her work.
* Designed Embedded Systems with Keil and microchip pic micro-controllers and exposed to various compilers and debuggers.
* Experience in configuration and requirement management tools such as DOORS and ClearCase.
* Good Exposure to C programming standards such as Code check and MISRA C.
* Exposure to Autosar Architecture.
* Worked in Robert Bosch for Parkpilot.

**Software Engineer,** *Artech Infosystems, Bangalore, India Aug 2016 – Feb 2017*

* Involved in generating test cases from the code taken from MKS Integrity tool.
* Reporting bugs and documenting.
* Working on C and C++ in both Windows and Linux Operating systems.
* Technical knowledge consists of Embedded Systems and serial Communication Protocols like UART, SPI, I2C, and UDS Protocols.
* Good Experience in various types of testing such as unit, regression, and Whitebox/Blackbox testing.
* High-performance coding in ARM Assembly, Python, Java, C++, C, Embedded C.
* I have hands-on experience writing scripts using Python scripts.

**Contract Assignee,** *Tata Consultancy Services, Hyderabad, India Aug 2016 – Feb 2017*

* Involved in developing and implementing test cases using the Cantata tool according to software requirement specifications provided by the customer.
* Found bugs and reported them to the responsible person.
* Provided the daily status of the work to the manager and got it approved.
* Designing, developing, and modifying unit tests based on customer requirements.
* Experience in working with multi-threaded Real-Time Operating Systems such as VXWorks.
* Proficient in Embedded C/C++ Programming and Assembly Programming on multiple platforms.

# ACADEMIC PROJECTS

**Multi-Label Clinical Text Eligibility Classification and Summarization System**

* Developed a system leveraging NLP and LLM models for automated multi-label classification and summarization of clinical text eligibility.
* Utilized word embedding (word2vec), entity recognition, count vectorizers, and TF-IDF for feature extraction.
* Implemented random forest and SVM models for multi-label classification based on eligibility criteria.
* Explored summarization techniques including TextRank, Luhn summarization, and GPT-3 for concise eligibility requirement summaries.
* Evaluated system effectiveness using ROUGE scores.

**Technologies Used: NLP, LLM models, word embedding (word2vec), TF-IDF, random forest, SVM, TextRank, Luhn summarization, GPT-3, ROUGE Score.**

**Predictive Modeling vs. Monte Carlo Simulation: Portfolio Optimization Analysis**

* Conducted comparative analysis of portfolio optimization strategies using predictive modeling and Monte Carlo Simulation.
* Leveraged advanced computational techniques like Monte Carlo Simulation and ML algorithms.
* Analyzed effectiveness of portfolio optimization strategies by comparing simulated and predicted returns with actual stock data.
* Utilized CAPM analysis for estimating expected return based on risk characteristics.
* Employed technical indicators for comprehensive analysis and prediction.
* Utilized Linear Regression and LSTM models for predicting stock prices.
* Employed mean-variance optimization and Sharpe ratio maximization for portfolio allocation strategies.
* Conducted distribution analysis of returns and assessed Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR) for risk assessment

**Technologies Used: Predictive modeling, Monte Carlo Simulation**

**Python Powered Stock Market Analytics with MongoDB, PySpark, and Tableau**

* Utilized Apache Spark, MongoDB, and PySpark for efficient management and analysis of large volumes of financial data.
* Analyzed historical stock market trends, cycles, and anomalies.
* Employed Tableau for creating interactive dashboards to present insights.
* Integrated free APIs provided by Twelve Data and Yahoo Finance for data retrieval.
* Leveraged PyMongo for seamless interaction with MongoDB.

**Technology and Tools used: Apache Spark, MongoDB, PySpark, Tableau, PyMongo, Twelve Data API, Yahoo Finance API**

**Predicting Patient Outcomes using ASA Classifications**

* Utilized machine learning techniques to predict patient outcomes based on ASA classifications.
* Conducted comprehensive data analysis and preprocessing of healthcare data.
* Implemented classification models to predict ASA classifications.
* Addressed class imbalance through random oversampling techniques.
* Evaluated model performance metrics including accuracy, precision, recall, and F1-score.
* Contributed to improving patient care, risk management, and surgical decision-making within the healthcare domain.
* Predicted discharge time (dis time) using regression techniques such as Random Forest Regression.
* Conducted feature engineering and selection to enhance prediction accuracy.
* Evaluated regression model performance using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).
* Incorporated the prediction of discharge time to provide a holistic understanding of patient outcomes and optimize healthcare resource management.

**Technology used: Machine Learning: Python, scikit-learn (sklearn); Data Analysis: Pandas, NumPy; Data Visualization: Matplotlib, seaborn; Web Development: Flask (for building web applications to visualize results or interact with models)**

**Financial Data Analysis:**

* Conducted data extraction, transformation, and loading (ETL) process.
* Utilized SSMS and SSIS for data management and ETL.
* Integrated transformed data into Power BI for reporting and visualization.

**Technology used - ETL, SQL, Power BI**

**Project Management System:**

* Designed classes using OOP principles for project management.
* Implemented classes in Python with defined attributes and methods.
* Applied efficient search and sorting algorithms.
* Conducted rigorous testing (unit, integration, system).
* Iteratively refined system based on testing feedback.
* Successfully deployed project management system.
* Emphasized thorough documentation for future maintenance.

**Technology used - Python, Object-Oriented Programming, Search Algorithms, Sorting Algorithms, Testing**

**London Visitor Analysis:**

* Cleaned the data by addressing missing and duplicate values and integrated various datasets by aggregating the data using Pandas and NumPy to extract the important information from a raw dataset of more than 60k records.
* Developed an interactive and dynamic dashboard using Python Dash, Pandas, NumPy, and HTML to answer the formulated hypotheses like whether people from countries with the highest GDP or Happiness index travel more, how happiness index, life expectancy, and GDP of a country are interlinked together by merging various datasets.

**Technology used – Python, Pandas, Dash, NumPy, HTML, Matplotlib**

# CERTIFICATIONS

* **CJET in Embedded Systems** *April 2015*
* **Foundations of Project Management Offered by Google issued by Course era** *August 2021*
* **HTML, CSS, and Javascript for Web Developers Offered by Johns Hopkins University** *July 2021*
* **Complete Java Certification Program Udemy** *Sep 2020*

# CONTACT INFORMATION

**Cell: +1 667-334-1464**

**Email: [almasfathima4@gmail.com](mailto:almasfathima4@gmail.com) / [fn40732@umbc.edu](mailto:fn40732@umbc.edu)**