**Yuvaraj Durairaj**

**Contact Information:**  
Email: [yuvaraj.d85@gmail.com](mailto:yuvaraj_durairaj@vanguard.com) | Mobile: +1 (312) 459-8667

### **Professional Summary**

Seasoned Application Engineering Technical Lead with extensive experience in Big Data ecosystems and AI/ML model development. Adept at leveraging industry-standard frameworks such as Apache Spark, Hadoop, PyTorch, Keras, and Scikit-learn to deliver scalable and intelligent solutions. Proven track record in full-cycle software development across diverse platforms including Java J2EE, Mainframes, and relational and distributed databases such as PostgreSQL, DB2, Oracle, MySQL, and Hive.

Highly proficient in multiple programming languages—Java, Scala, Python, and COBOL—bringing versatility and precision to complex engineering challenges. Demonstrated expertise in designing and deploying machine learning models tailored to business needs.

Over 10 years of hands-on experience architecting and managing cloud-native applications on AWS, with deep knowledge of services including S3, ECS, EC2, CloudFormation, CloudWatch, Glue, EMR, Lambda, IAM, KMS, Step Functions, DynamoDB, Aurora RDS-Postgres, Route 53, Bedrock, and SageMaker.

### **Core Skills**

* **AI/ML Tech Stack**: PyTorch, SciKit, Keras, SageMaker
* **Big Data Technologies**: Apache Hadoop, Apache Spark, Hive, Sqoop, Oozie
* **Cloud Platforms**: AWS (ECS, Glue, Lambda, EMR, S3, DynamoDB, Athena)
* **Programming Languages**: Java, Scala, Python (Polars, Pandas, NumPy, Flask, Dependency-Injection, Scikit, Matplotlib)
* **Frameworks**: Spring Boot,RESTful Web Services - Jersey & Spring MVC, TensorFlow, Keras, PyTorch,Hugging Face & Hibernate
* **Databases**: PostgreSQL, Oracle, DB2, Apache Hive
* **Development Tools**: Maven, Jenkins, Bitbucket, GitHub, Bamboo, Cucumber
* **Other Skills**: Shell Scripting, Data Analytics

### **Professional Experience**

#### **The Vanguard Group, Inc.**

**Application Engineering Tech Lead III**  
*Malvern, PA (Mar 2021 – Present)* - Led modernization of APIs for EIP performance metrics (Returns, Fees, Yields, Expense Ratios) and refactored data ingestion to decommission mainframes. - Developed scalable solutions for internal and external investment product data dissemination. - Implemented cost-efficient data ingestion frameworks, reducing operational costs by 90%. - Have developed AI Models for DQDL Rules generation. - Have developed RAG based model for hackathon using Bedrock for SQL generation using semantic natural language.

Environment/Tech Stack: Spring Boot, AWS (ECS, Glue, Lambda, RDS), Apache Spark (Scala & PySpark)

**Application Engineering Tech Lead II**  
*Malvern, PA (Apr 2019 – Mar 2021)* - Designed AWS-based data lakes for corporate finance analytics, enabling real-time insights. - Built smart data ingestion tools exceeding AWS Glue Crawler’s functionality, supporting large-scale data. - Created a data lake architecture to centralize and optimize internal audit operations. - Experimented Vanguards Opex savings projection using Regression model.

Environment/Tech Stack: Apache Spark, AWS (EMR, Glue, S3, Lambda), Oracle DB

**Application Engineering Developer III**  
*Malvern, PA (Apr 2017 – Mar 2019)* - Engineered compliance reporting workflows (Class Actions, Corporate Actions) using AWS and Spark. - Collaborated with RDnA core team to load enterprise datasets into AWS data lakes for analytics.

Environment/Tech Stack: Python, Apache Spark, AWS (Glue, EMR, S3)

#### **Randstad Technologies (Client: Vanguard Investments)**

**Software Developer**  
*Malvern, PA (Jan 2016 – Apr 2017)* - Built Big Data frameworks for generating retail investment insights using Apache Spark and AWS. - Developed RESTful APIs for data pipelines, facilitating seamless cloud data transfer.

Environment/Tech Stack: Java, Spark, RESTful Services, AWS (Glue, EMR, S3)

#### **Tata Consultancy Services**

**Module Lead/ Senior Developer (Client: Bank of America)**  
*Addison, TX (Nov 2012 – Jul 2013)* - Designed and implemented Restful services for large-scale financial reporting. - Coordinated architectural design and streamlined data integration processes.

Environment/Tech Stack: Java, Spring Boot, AngularJS, DB2

**Business Analyst (Client: Bank of Montreal / Harris Bank)** *Chicago, TX (Feb 2012 - Nov 2012)* - Analyzed applications across verticals (Deposits, Corporate & Shared Services, Treasury, Lending, and Channels) for the M&I (source) to BMO-Harris (target) Conversion project.  
- Conducted system gap analysis and managed data mapping and data cleansing activities.  
- Contributed to the Target Operating Model team, merging the best features from both systems into a unified platform for improved operations.

**Module Lead/ Senior Developer (Client: Morgan Stanley)**  
*Chennai, India (Feb 2010 - Nov 2011)* - *Smith Barney to Morgan Stanley Data Conversion (Jul 2010 – Nov 2011):*

- Gathered requirements for data conversion and developed reusable components for multiple modules.  
- Lead the Trick Account data conversion team, addressing complex business requirements and customer data intricacies.  
- Designed and implemented a one-time application to manage complex data conversion tasks.

* *Branch Number Expansion – Strategic Solution (Feb 2010 – Jun 2010):*
  + Developed and tested solutions to integrate Smith Barney branches into the Morgan Stanley system.
  + Conducted end-to-end system testing and provided support for seamless implementation.
  + Analyzed specifications to address system modifications and joint venture program requirements.

#### **Satyam Computer Services Ltd.**

**Software Developer (Client: Nissan North America)** *Chennai, India (Aug 2006 - Oct 2009)* - Supported maintenance and production processes across six business domains: Corporate Functions, Sales & Marketing, Manufacturing, Auto Finance, Engineering, and Nissan Canada (NCI).  
- Worked on the critical business application for Production Ordering & Scheduling, handling production support, enhancements, development, testing, and implementation.  
- Quickly grasped business functionality and assumed full ownership of the application, ensuring 100% compliance.  
—

### **Education**

* **MBA (Finance and Investment)** – Drexel University - Lebow School of Business (Class of 2025)
* **B.E. in Mechanical Engineering** – Anna University - St. Peter’s Engineering College (Class of 2006)

### **Certifications**

* AWS Developer Associate Certification
* AWS AI Practitioner

### **Achievements**

* **Spot Award (Vanguard):** For scalable architecture design supporting internal tools.
* **Star of the Sprint (Vanguard):** Recognized for resolving critical defects in corporate action systems.
* **Pat on the Back Award (Satyam):** Delivered zero-defect production releases for major development projects.

### **Technical Proficiencies**

* **Programming:** Java, Scala, Python, COBOL, Natural, Shell Scripting
* **Big Data:** Apache Hadoop, Spark, Hive, Oozie, Flume
* **Cloud Platforms:** AWS (ECS, Glue, Lambda, EMR, S3, SageMaker, StepFunction)
* **Databases:** PostgreSQL, Oracle, DB2, ADABAS
* **Frameworks:** Spring Boot, PyTorch, LangChain

### **Notable Projects**

* **Social Sentiment Analysis:** Developed Big Data pipelines to analyze mutual fund sentiment using Twitter and Facebook APIs.
* **Transactional Analytics App:** Built a web application for segmenting consumer expenditures using Big Data and Java.
* **Reusable Tools for Vanguard CTO Team:** Contributed scalable solutions benefiting enterprise-wide AWS adoption.
* **Query Builder Service:** Designed and implemented a dynamic query generation service that enables users to construct SQL queries through intuitive selection inputs, eliminating the need for manual coding. Enhanced the solution by integrating AWS Bedrock with foundational models (Claude) to interpret business semantics via prompt engineering. Leveraged LangChain to maintain contextual continuity, enabling accurate and adaptive SQL generation aligned with enterprise data structures.
* **Data Valiation Service:** Engineered a lightweight alternative to AWS DQDL for dataset validation based on configurable rule sets. Utilized the Polars framework to efficiently execute validation logic and dynamically capture triggered rules as dedicated output columns, enabling streamlined data quality assessments and enhanced traceability.