## Subba Reddy Yeruva

[subbareddyyeruvaus@gmail.com](mailto:subbareddyyeruvaus@gmail.com) | + 1 646 744 7826

www.linkedin.com/in/subbareddyyeruva | <https://github.com/SubbaReddyYeruva> | https://rpubs.com/yeruva\_s\_reddy

Versatile Professional: Offering over 10 years of rich experience with core strengths in Banking industry:

**Data Science** | **Big Data** | **Product Development**

“High interest and core competencies in Data Science”

|  |  |
| --- | --- |
| KEY SKILLS  Product Development  Business Analytics  Data Science  Python, R, Tableau  Statistics  Regression Modelling  Data Mining  Machine Learning  Data Wrangling  Big Data  Spark  SCALA  NoSQL  Retail Banking  Sales and Marketing    Performance Engineering  Mainframes | **PROFILE SUMMARY**   * Performed extensive data studies to understand the statistical behaviors of the upstream machines * Designed and implemented data ingestion techniques for real time data coming from various data sources * Built predictive analytics models to generate actionable insights * Implemented clustering and segmentation techniques * Performed business analysis by defining the requirements. Designed BRD, Functional specifications * Mentored teams   .   * Key **Analytics** Competencies: * Statistical Analysis: Estimation and Testing, Regression modelling and advanced methods * Operations: Simulations and Optimization * Data Mining: Supervised and Unsupervised learning * Forecasting Analytics: Time Series Modelling, ARIMA and etc. * Contemporary Analytics: Strategies, Digital and Social Media, Analytics for new product development. * Data Collection and exploration (Python, R) + Data Visualization * Financial Analytics * Big Data and NoSQL Expertise * Natural Language Processing (NLP) * Possess strong analytical and problem solving skills; effective leader with excellent skills in motivating individual employee performance |

TECHNICAL SKILL-SET

* **Analytical Tools**

SQL, R-Studio, Jupyter Notebook, Tableau, Dataiku, XLMiner, NodeXL, STATA, @Risk, Graph Database, H2O, AtScale

* **Programming**

R, Python, SCALA, Cypher

R - Data Manipulation, Analytical Packages, Shiny R

Python - Data Manipulation, Numpy, Pandas, Matplotlib, Plotly, Scikit learn (machine learning libraries and others)

* **Big Data**

Spark, Hive, Kafka, Pig, Apache Drill, Dremio, Flume, Sqoop, HBase, Impala, Hadoop, HDFS, YARN, MapReduce

Spark - Spark Core, Spark SQL, Spark Streaming, MLlib, GraphX, PySpark, Scala, Data Frames, and SparkR

* **NoSQL**

Neo4j, MongoDB, Cassandra

* Others

AWS, Mainframe, DB2, Cognos, SQL Tunining, Java Basics, SAS E-Miner

EDUCATION

Business Analytics (Data Science) - Indian School of Business(ISB), Hyderabad, India

B.E. (Electronics and Communication Engineering) - Andhra University, Vishakhapatnam, AP, India

Key Result Areas

* Data Analysis
* Frameworks development in Big Data catering to Analytics
* Built predictive analytics models for machines to generate actionable insights to minimize the down time
* Requirement analysis and Gap analysis
* Preparing functional design and technical design
* Coordinating with business stakeholders to understand business problems and opportunities in the context of requirements and suggesting solutions to achieve stakeholder’s goals
* Implemented productivity initiatives by developing internal process and automating the process

Accomplishments

#### Best Business Facilitator Award: For assisting the training department to deliver trainings and mentoring the freshers

* OHCB Star Award: For involving in multiple domains proactively and reduced the defect ratio.
* TEAM Awards: Multiple team awards, for collaborating with multiple departments and delivering high quality product

Projects

* **Morgan Stanley - Machine Learning Engineer – Analytics AI and Data Tech - Jan, 2019 to Current**

Company : Morgan Stanley

Division : Analytics AI and Data Tech

Department : Wealth Management Analytics and Reporting

Key Responsibilities:

* Building distributed, high-volume data pipelines to power new product features based on analytics and machine-learning needs
* Implementing machine learning approaches to improve customer experience (recommendation, similar products, etc.)
* Making the business to adopt the Big Data Analytics offered features
* Building the Big Data applications which will cater to high velocity, volume and variety of data
* Building the data visualizations for regulatory purpose
* Building the recommendation system for campaign management by leveraging Big Data and Data Science
* Productionizing the Big Data and Data Science applications
* Supporting the Big Data and Data Science applications
* Carrying out data analysis/studies to fix the production issues
* Building applications for Machine Learning, Natural Language Processing (NLP)/Text mining
* Involving from requirements elicitation to end product delivery and communicating the results to stakeholders
* Evaluating the performance of Machine Learning applications
* Developing Realtime scoring models
* Involving from design to implementation of machine learning models
* Defining the requirements for data lakes/pipe lines
* Automating the scripts using Python
* Developing data mining applications using Spark
* Transforming or extracting the features using Spark for semi structured data
* Performing extensive data studies to understand the statistical behaviors
* Design and implement data ingestion techniques for real time data coming from various source systems
* Defining the data layouts and rules and after consultation with ETL teams
* Designing technical and business process flow for the analytics projects
* Developing the data pipelines, performance optimization of data pipelines and machine learning applications
* Carrying out Proof of concepts on new tools and technologies to adopt in Big Data and Data Science applications
* **United Health Group - Data Scientist – Advanced Analytics Lab - May, 2017 to December 2018**

Company : United Health Group

Division : Product Engineering and Data Solutions

Department : Advanced Analytics Lab

Key Responsibilities:

* Analyzing the claims data for over payments and leakage of claims. Building the predictive models and cost analysis
* Developing Realtime scoring models
* Involved from design to implementation of machine learning models
* Defining the requirements for data lakes/pipe lines
* Perform extensive studies of different technologies and capture metrics by running different algorithms
* Converting the SAS algorithms into different technologies like PySpark
* Automated the scripts using Python
* Developed data mining applications using PySpark
* Along with other data ingestion tools used PySpark for data ingestion and data transformations
* Transforming or extracting the features using PySpark for semi structured data
* Perform an extensive data studies to understand the statistical behaviors
* Applied machine learning algorithms to identify the overpayments
* Design and implement data ingestion techniques for real time data coming from various source systems
* Defining the data layouts and rules and after consultation with ETL teams
* Designed technical and business process flow for the upcoming advance analytics project
* Conducting trainings to different teams to adopt Big Data technologies
* Defining the metrics for the Big data analytics proof of concepts
* Leading the projects to achieve key business objectives
* Handled the implementation activities and provided Post implementation support

Achievements:

* Developed predictive model to identify the Over Payment of claims
* Assisted business by reducing elapsed time of applications and hence time to market is reduced by streaming the applications
* **HSBC - Data Scientist – Sales and Marketing System - Jan, 2016 to May, 2017**

Company : HSBC (Hongkong and Shanghai Banking Corporation)

Division : Product and Engineering Services

Department : GIS (Global Information Services)

Key Responsibilities:

* Data analysis using open source tools
* Defining the requirements for data lakes/pipe lines
* Perform an extensive data studies to understand the statistical behaviors
* Predicted the application score and the behavioral score of the customers by applying ensemble methods (Regression methods and random forests)
* Developing predictive models
* Applied machine learning algorithms
* Developed the applications in PySpark framework
* Design and implement data ingestion techniques for real time data coming from various source systems
* Creation of regulatory reports and analysis. Defining the data streams
* Defining the data layouts and rules and after consultation with ETL teams
* Built predictive models for cross selling the products
* Sales forecasting was done using time series forecasting
* Designed technical and business process flow for the coming advance analytics project

Achievements:

* Created the 360-degree view of the customers as per digital strategy
* Created predictive models for ATM locators and cash flows
* Complaints handling process were streamlined
* Campaign management process was streamlined
* Helped business by creating revenues (Data Monetization) with 30% of conversion
* **HSBC - Data Scientist - HCB (HSBC Core Banking) -May, 2013 to Dec, 2015**

Company : HSBC (Hongkong and Shanghai Banking Corporation)

Division : Product and Engineering Services

Department : RBWM (Retail Banking and Wealth Management)

Key Responsibilities:

* Data analysis using open source tools
* Defining the requirements for data lakes/pipe lines
* Automated the scripts using Python
* Forecasted quality issues and defined cautionary measures
* Applied clustering and segmentation methods for product offerings
* Reports creation and analysis. Examples include Currency level report(CLR), Account Level Report(ALR) and other reports.
* Defining the data streams
* Interacting with all the stakeholders and defining and making sure that the data flows are fine

Achievements:

* The critical gaps in the system were fixed
* **HSBC - Business Analyst - HCB (HSBC Core Banking) -Feb, 2012 to Apr, 2013**

Company : HSBC (Hongkong and Shanghai Banking Corporation)

Division : Product and Engineering Services

Department : RBWM (Retail Banking and Wealth Management)

Key Responsibilities:

* Preparing BRD and traceability matrix
* Preparing Functional Design Documents based on Gap Elaboration Document and BRD.
* Defining the functional test cases and functional flows.
* Review of test cases
* Performing UAT
* Review of functional design documents

Achievements:

* Delivered the project with zero defects
* **HSBC - Senior Software Engineer - HCB (HSBC Core Banking) - Oct, 2009 to Jan, 2012**

Company : HSBC (Hongkong and Shanghai Banking Corporation)

Division : Product and Engineering Services

Department : RBWM (Retail Banking and Wealth Management)

Release 1.0 : HCB - Saving Accounts (October 2009 to July 2010)

Release 2.0 : HCB - Term Deposits (August 2010 to January 2011)

Release 3.0 : HCB - Current Accounts (February 2011 to July 2011)

Release 3.1 : HCB – Package and Foreign Exchange Accounts (August 2011 to January 2012)

**Description**: The Objective of the overall HSBC Core Banking program is to streamline HSBC’s business processes and products around the needs of its customers on a global basis.

Key goals of the program are:

* + Enhance customer experience to fully reflect the promise of HSBC as a brand.
  + Accelerate cost savings from standardization of technologies, products and processes on global scale.
  + Facilitate delivery of global propositions, leveraging the group’s unrivalled distribution.

**Key Responsibilities as Part of development**:

* Preparing Functional Design Documents based on Gap Elaboration Document and BRD.
* Preparing Technical Design Documents based on Functional Design.
* Preparing Technical Specification Documents based on Technical Design.
* Coding of Service layer programs.
* Involving in Unit Test Plan and Unit Test Results.
* Involving in FD, TD, TS, Code, UTP, UTR Reviews.
* Involved in UIT, AIT and SIT phases as part of Development Team.
* Involved in OAT Phase as part of deployment support.

**Key Responsibilities as Part of Performance Tuning**:

* Preparing the test strategy for Online and Batch Application testing.
* Preparing the data as per requirement (1 Million to 5 Million Accounts).
* Pre-assessment of the application before Performance tuning changes.
* Analyzing the APA and CICS Trace.
* Reviewing the application for tuning opportunities.
* Preparing the CCJD (Code Change Justification Document) for the identified application changes.
* Final Assessment of the application after Performance tuning changes.

* **Freelance - Software Engineer - Technology - Apr, 2008 to Sep, 2009**

Key Responsibilities as a freelance:

* Gathering and eliciting the requirements
* Preparing the technical design documents