

# Shyam BV

Linkedin/Medium/Github/Website

Email : shyambv86@gmail.com

Mobile : +1-908-838-8800

## EDUCATION

---

- **City University of New York** New York, NY  
*Master in Data Science; GPA: 4.00*
- **Anna University** Chennai, India  
*Bachelor of Engineering in Electrical and Electronics; GPA: 3.43*
- **National Institute of Information Technology** Trichy, India  
*Diploma in Information Technology;*

## EXPERIENCE

---

- **Equifax - Data Scientist** Atlanta, GA  
*Customer facing Data Scientist* *Sep 2018 - Present*
  - **Identity and Fraud - Attributes and Models:** Creating fraud attributes using PII's and models using the attributes.
  - **Identity and Fraud - Graph:** Created graph attributes to find fraudulent accounts using PII's in Equifax and execute it large scale.
  - **Ignite Direct - Models:** Helping Equifax customers to solve their analytics use cases with various Equifax data assets. Created models on propensity to buy and scoring for customers.
- **Multidimensions - Self-employment** Atlanta, GA  
*Solopreneur* *Feb 2020 - Present*
  - **Code Sprout:** A medium and substack blog to publish different articles on applications of code.
  - **Socialimageai.com:** AI marketing product for startups to create social media posts, articles from website or context using GPT-2, image recognition
  - **Sproutquik.xyz - Retired:** Created a product for startups to get up and run quickly. Similar to AWS amplify. Platform is built using terraform and AWS.
  - **Buildtheresume.com:** Product to create professional resume using OpenAI GPT-3 to create resume and perform resume QA using sentiment and transformer models.
  - **tradeslybot.com - Retired:** Customer facing chatbot via social media channels(Whatsapp, FB, Telegram) to trade and provide information on Stocks.
- **Deutsche Bank - Data Scientist** Cary, NC  
*Assistant Vice President* *Sep 2014 - Sep 2018*
  - **Fraud Detection:** Replacing the existing rule-based system by analytics application to perform anomaly and fraud detection in the retail banking. Created models(parametric/non-parametric) and architect the complete pipeline for the real-time application.
  - **Texas Oil Estimation:** Created seasonal estimation models using ARIMA/Prophet/Neural Networks to predict the output in Texas oil wells and provide recommendations.
  - **COE - QA system:** Created QA(Chatbot) using Deep learning toolkit to address the common questions in Center of Excellence team.
  - **Internal Documents classification:** Classified Deutsche Bank internal document classification from Enterprise content management system using ensemble models and Neural networks.
  - **Credit/Operational Risk:** Developed credit/operational risk-based models and building a data pipeline using Cloudera data science workbench. Also, migrated existing SAS based risk models to pyspark.
- **Tata Consultancy Services - Sr. Data Analyst** Edison, NJ  
*AIG - Consultant* *Mar 2011 - Sep 2014*
  - **ISG Cost Containment:** Performed various data analysis/mining activities and expenses by applications. Created models to predict the risk and expected over-expenses.
  - **DWFIT:** Worked as a Data Engineer to build data pipelines for an enterprise-wide Financial Information Transparency Data warehousing system on Hadoop.
  - **Financial Information Transparency Budgeting:** Performed data analyst role by creating different forecast/budgeting cost-based models.

- **Sensiple Softwares - System engineer**

*Data Engineer*

*Jul 2008 - Mar 2011*

- **West Connect:** Performed analysis of voice metrics analysis and created various KPI's and metrics with Business intelligence tools.
- **Mileage Plus:** Performed ETL transformations on the mileage plus voice application from transactional to OLAP database.
- **Katana:** Worked on an OTC bond market trading application using core Java. It is a real-time trading application for banks and brokers.

## TECHNICAL SKILLS

---

- **Programming/Scripting:** Python, Java Script, Web Development, PLSQL
- **ML:** Pytorch, Computer Vision, NLP, Machine Learning, Deep Learning, SparkML, H2O
- **Data Engineering/DevOps:** Terraform, Hadoop, PySpark, Docker, Cloudera Data Science Workbench, CICD, AWS, GCP
- **Visualizations/Databases:** Tableau, Qlikview, BO, Cognos, Neo4j, Oracle, AWS DynamoDB
- **Stat/Math/Others:** Statistical analysis, Regression, A/B tests
- **Management:** Founded companies called 'Stainwash'(Retired) and 'Multidimensions'