

CIJO GEORGE

Email: cijogeorge.in@gmail.com, **Phone:** +91 99 4567 0410

LinkedIn.com/in/cijogeorge, **GitHub.com**/cijogeorge

PROFILE SUMMARY

Data Scientist/ Machine Learning Engineer with interest and experience in statistical analysis and machine learning techniques, frameworks and platforms. Passionate about designing and developing end-to-end data-driven solutions to challenging problems using state-of-the-art/ emerging technologies.

TECHNICAL SKILLS

- **Statistical analysis and machine learning:** Time-series analysis, forecasting and anomaly detection, classification, clustering, feature engineering, dimensionality reduction, model validation and evaluation techniques, intuitive understanding of mathematics behind learning models.
- **Programming languages:** Python (scikit-learn, pandas, flask et al.), R (caret, dplyr, forecast, shiny et al.), BASH scripting (sed, awk et al.), SQL, C, C++, MPI, CUDA, Perl, PHP, HTML
- **Frameworks, platforms and tools:** Hadoop (Hive, MapReduce), Jupyter, Docker, Nginx, Gunicorn, Git, Amazon Web Services (AWS), Cask Data Application Platform (CDAP), LaTeX.
- **Other skills:** RESTful API design, agile development, working knowledge of operating systems, storage systems and networks.

PROFESSIONAL EXPERIENCE

NetApp

August 2012 - Present

Member Technical Staff/ Data Scientist, Advanced Technology Group

Bangalore

- Lead advanced development projects with responsibility of studying the problem, proposing a solution, developing a working prototype/ proof-of-concept and taking prototypes into production.
- Successfully completed multiple big data analytics projects involving statistical analysis and machine learning techniques including *time-series analysis/ anomaly detection/ forecasting, classification, clustering, feature engineering* and *dimensionality reduction* on machine generated data.
- Designed and developed software to deploy statistical and machine learning models in production, with scalable *RESTful API interfaces* and *scale-out processing* of data analysis jobs for specific use-cases.
- Designed and prototyped a *data virtualization* solution providing a *data-pipeline* based approach for community-driven development of a semantic data catalog, with the goal of absorbing common machine learning and advanced analytics based data transformations into the data platform layer.
- Co-authored technical reports on emerging technologies in big data analytics/ machine learning.
- Co-authored an external publication on identifying/ troubleshooting performance problems in storage systems based on anomaly detection on time-series metrics collected from live systems.
- Involved in pitching new project ideas, interacting with stakeholders (technical and non-technical), mentoring juniors/ interns and collaborating with researchers from top Indian universities.

Nokia Siemens Networks

June 2009 - July 2010

Software Engineer, R&D

Bangalore

- Part of development team for HLR (Home Location Register), a key component for call handling and value added services in the telecommunications *core network*.
- Involved in feature development for the next generation HLR release versions as well as maintenance of older versions.

EDUCATION AND RESEARCH EXPERIENCE

Indian Institute of Science, Bangalore

June 2013

Supercomputer Education and Research Centre

Degree: Master of Science in Engineering (by research)

Thesis Title: Fault Tolerance Strategies for Large Scale Systems

- Developed and prototyped a fault avoidance technique for MPI applications based on partial replication of processes and dynamic changing of replica set, using failure probability models for predicting impending failures.
- Developed a fault avoidance technique for malleable applications that dynamically changes the fault tolerance mechanism of the application throughout its execution, using failure probability models for predicting impending failures.
- Developed a generic fault tolerance simulator for parallel computing systems.
- *Course project:* Developed an adaptive technique to maximize resource utilization of physical systems in a Cloud environment based on resource utilization metrics.
- *Course project:* Worked on implementation of thread scheduling, system call interface, virtual memory and file system in Pintos OS framework.
- *Teaching assistantship:* Worked as a TA for two graduate level courses - High Performance Computing, Parallel Programming.

Cochin University of Science and Technology

April 2009

Government Model Engineering College

Degree: Bachelor of Technology, Computer Science and Engineering

- *Course projects:* USB mass storage device sharing over TCP/IP network; A collaborative text editor.
- *Pet projects:* An online stock-exchange simulator with real-time data feeds from National Stock Exchange; BASH-as-a-service: A web-browser based remote interface to BASH.

EXTERNAL PUBLICATION

- **Cijo George**, Sathish Vadhiyar, “Fault Tolerance on Large Scale Systems using Adaptive Process Replication”, *IEEE Transactions on Computers*, September 2014.
- Vipul Mathur, **Cijo George**, Jayanta Basak, “Anode: Empirical Detection of Performance Problems in Storage Systems using Time-Series Analysis of Periodic Measurements”, *MSST '14: Proceedings of the 30th International Conference on Massive Storage Systems and Technology (IEEE)*, Santa Clara, California, June 2014.
- **Cijo George**, Sathish Vadhiyar, “AdFT: An Adaptive Framework for Fault Tolerance on Large Scale Systems using Application Malleability”, *ICCS '12: Proceedings of the International Conference on Computational Science (Elsevier)*, Omaha, Nebraska, June 2012.

PATENTS FILED

- Vipul Mathur, **Cijo George**, Swaminathan Ramany, “System and Method for Analyzing a Storage System for Performance Problems using Parametric Data”, United States 14/194,467, Filed: February 28, 2014.

CONTACT DETAILS AND SOCIAL PROFILES

- **Email:** cijogeorge.in@gmail.com, **Phone:** +91 99 4567 0410
- **LinkedIn.com**/in/cijogeorge, **GitHub.com**/cijogeorge