

Russi Chatterjee

- Github: [@ixaxaar](#)
- Stackoverflow: [@ixaxaar](#)
- Twitter: [@markovinchains](#)
- Email: root@ixaxaar.in / russichatterjee@gmail.com

Short

- Data Scientist and Engineer with ~ 8 years exp.
- Experience in Ad-tech, healthcare, linguistics, social networks, IOT and embedded systems.
- Blend of data science and engineering experience.
- Held leadership roles for > 4 years.

Stack

- **Languages:** Scala, Python, bash
- **ML toolkit:** scikit-learn, scikit-image, networkx, pytorch, pymc3, pandas, numpy, scipy, CoreNLP, KenLM, openIE, ML-LIB
- **Big data:** Spark, presto, hive, storm, AWS EMR, Athena, Glue catalog
- **Databases:** cassandra, elasticsearch, mongo, postgres / postGIS
- **Others:** jupyter, atlassian suite, devops tools, etc

Experience

1. Senior Data Scientist, [Lifesight.io](#)

[Geo-location ad analytics \(05-2018 - present\)](#)

- Lead the data science team.
- Worked on various ML models on classification and clustering.
- Worked on data engineering jobs for geospatial problems.
- Architected various components of the big data stack.
- Planning, tracking, brainstorming and managing tasks.
- Processed petabyte-scale data.

2. Data Scientist, **Reverie**

Creating an NLP stack for Indian languages (07-2016 - 08-2017)

- Individual contributor role amongst a team of researchers.
- Worked on:
 1. cross-linguistic information retrieval
 2. neural transliteration
 3. language models
 4. queryable knowledgebase

3. Data Scientist, **Practo**

India's largest healthcare platform (05-2015 - 07-2016)

- Lead the search team.
- Conceptualized, architected and executed "Cerebro", the search engine that powers practo.com.
- Authored the doctor ranking algorithm
- Involved in various data science tasks across verticals.

4. Senior Software Engineer, on contract, **Showt**

Real-time public opinion network (07-2014 - 01-2015)

- Data engineering role.
- Involved designing distributed systems at a scale of 20k QPS.

5. Senior Software Engineer, MagnetWorks

Analytics and monitoring for process industries (08-2013 - 06-2014)

- IOT analytics product for heavy industries.
- Extremely hands-on, full-stack developer role.

6. Software Engineer, Toshiba

Real-time operating systems (02-2013 - 08-2013)

- Worked with the microkernel core OS team, Japan.

7. Software Engineer, Wipro

Automotive Electronics (02-2011 - 02-2013)

- Linux kernel device driver developer.

Opensource

- **Deep Learning**

- [pytorch-dnc](#) : Differential Neural Computers, Sparse Access Memory and Sparse Differentiable Neural Computers.
- [pytorch-dni](#) : Implementation of Deepmind's "Decoupled Neural Interfaces Using Synthetic Gradients".
- [subLSTM](#) : Implementation of subLSTMs from the paper "Cortical microcircuits as gated-recurrent neural networks".

- **Computational Neuroscience**

- [Spectrum](#) : Higher-order spectrum estimation toolkit, for analysing EEG recordings
- [Cortical Microcircuit](#) : Implementation of a model of a cortical microcircuit (a 1mm² of the human brain), in pyNEST, from the paper "The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model."
- [S-transform](#) : The Stockwell transform.

- **Mathematics**

[monoid.space](#), a book on learning how to "implement" a few math theories.

Education

- B-Tech., Computer Science, [NIT, Rourkela](#) - 2006-2010