Sreejith Sreekumar

#1, 1 Folsome Ave, Roxbury Crossing, Boston, Massachusetts, 02120

857-399-6443 | sreekumar.s@husky.neu.edu | srjit.github.io | srjit | sreejith2904

Education

Northeastern University, Boston, MA   *Jan. 2017 - Present*

CANDIDATE FOR MASTER OF SCIENCE IN DATA SCIENCE   *Expected Graduation: Dec 2018*

- Related Courses : Supervised Machine Learning, Natural Language Processing, Applied Probability and Stochastic Processes, Special Topics in Artificial Intelligence

Government Engineering College, Thrissur   *Sep. 2007 - June 2011*

BACHELOR OF TECHNOLOGY

- Related Courses : Data Structures and Algorithms, Database Management Systems, Programming Paradigms, Numerical Analysis and Optimization Techniques, Design and Analysis of Algorithms

Technical Knowledge

Specialities	Classification and Clustering, Regression, Deep Learning, Natural Language Processing & Distributed Computing
Programming	Python, R, Scala, Shell Scripting, Java, Groovy, Javascript
Big Data Ecosystem	Apache Spark and Spark Mllib, Apache Hadoop, Hive, Flume, Sqoop, Oozie
Databases	MySQL, MongoDB, HP Vertica, H2
Domains	E-commerce
Certifications	Scalable Machine Learning(edX), Introduction to Big Data with Apache Spark (edX), Machine Learning (Coursera)

Academic Projects

- **Quantifying Semantic Similarity of Sentences using Long Short-Term Memory Neural Nets:** Using vectorized representations from Glove embeddings, the deep LSTM network designed could classify semantically similar questions from the ones which are not. The network was trained on a NVIDIA GeForce GTX 950 GPU and an accuracy of 83% was obtained.
- **Finetune AlexNet for Custom Classification :**
- **The Fake News Stance Classification:** Achieved a maximum accuracy of 88% on the task of classifying fake news from the genuine ones to four discrete levels - agree, discuss, disagree, unrelated using handcrafted linguistic features along with distance features from vectorized fields(Word2Vec). Random Forests, Support Vector Machines and XGBoost algorithms were used for performance comparison.

Experience

[24]7 Innovation Labs   *Bangalore India*

SENIOR DATA ENGINEER   *June 2016 - Dec 2016*

- Designed and developed a natural language toolkit on PySpark for chat transcript data analysis and modeling.
- Achieved an accuracy of 90% in modeling chat transcripts and classifying them for user intent prediction.
- Configured and scaled the toolkit on a multi-cluster environment with three apache spark nodes.

[24]7 Innovation Labs   *Bangalore, India*

DATA ENGINEER   *May 2015 - June 2016*

- Analyzed and modeled user data from web for several clients in the e-commerce domain for increasing chat propensity and uplifting purchases.
- Integrated SVM algorithm into the domain specific custom modeling tool and scaled over a million data points.

Xurmo Technologies Pvt. Ltd.   *Bangalore, India*

SOFTWARE ENGINEER   *July 2011 - May 2015*

- Developed and maintained machine learning modules of the flagship product of the company - Xurmo big data analytics platform.
- Developed and integrated machine learning algorithms on Apache Spark (Java).
- Developed HTTP based APIs for data retrieval and processing for the platform.
- Developed custom analytics functions for the engine data transformation.
- Wired up data ingestion and services APIs of the product.
- Developed analytics applications on the platform - Text exploration engine, Stock market movement prediction, Sentiment analyzer, Customer churn prediction.
- Implemented data collection modules for an enterprise knowledge discovery application.
- Collaborated for building data indexing and query optimization modules.