Tuhin Sharma

https://github.com/tuhinsharma121   tuhinsharma121@gmail.com   +91-886-155-7799

Career Profile

- 4+ years of core analytics/data science experience in helping businesses leverage data.
- Expertise in Machine Learning, Deep Learning and Natural Language Processing.
- Author @O'reilly Media for Deep Learning articles.

Technical Skills

- Languages : C,C++,Python
- Tools : Spark, Flink, Kafka, Zookeeper, Mongo, Druid, Scikit-Learn, Pandas, Numpy, Scipy, Pomegranate, Opencv, Tensorflow, Mxnet.

Work Experience

Co-Founder and CTO, Binaize, Bangalore	Dec 2018 - Present
Senior Data Scientist, Red Hat, Bangalore	Feb 2017 – Sep 2017
Data Scientist, IBM Watson, Bangalore	July 2014 – Feb 2017

Education

Indian Institute of Technology Roorkee (IITR)   **2012-2014**

M.Tech in Computer Science and Engineering, Data Mining Lab with CGPA 9.29

Thesis: A Generalized Framework for Quantifying Trust of Social Media Text Documents

Indian Institute of Engineering Science and Technology, Shibpur (IIESTS)   **2008-2012**

B.E. in Computer Science and Technology with aggregate 80% marks

Thesis: Dynamic Network Traffic Data Classification for Intrusion Detection Using Genetic Algorithm

Recent/Relevant Projects Details

- ❖ **Fabric-8-analytics (RedHat)** Provides recommendation for the manifest file of a developer in terms of alternate package, companion package and outlier package. Built Probabilistic Graphical Model to solve this problem. Code is available at https://github.com/fabric8-analytics/fabric8-analytics-stack-analysis
- ❖ **Demand Insight (IBM Watson)** Provides interesting insights about demands of products to be sold based on weather and social information about respective store locations. Built models in python using big data tool Spark.

- ❖ **Predictive Customer Analytics (IBM Watson)** Provides interesting insights in customer domain like Churn, CLTV, Segmentation etc. Built such models in python leveraging Spark ML/MLLIB.
- ❖ **Social Media Analytics (IBM Watson)** An automated system that provides interesting insights in financial domain by interpreting twitter data. The insights could be used for choosing different domain for doing sentiment or segmentation analysis and for targeting potential customers corresponding to individual financial banks.
- ❖ **FreeTime Designer (IBM Watson)** Generating a personalized optimal itinerary for users using events in their city to get engaged in their free time.
- ❖ **Twitterati Speaks (IBM Watson)** Finding top events related to a cricket match using social media using ngram analysis followed by hierarchical clustering.

Recent Mentoring/Consulting (not exhaustive)

Recommender System for News Media (Recotap)   May 2018 – Dec 2018
Provides generalized and personalized recommendation which can be fine-tuned effectively and efficiently based on different types of customer and market segments.

Digital Worker (Ivanti)   Nov 2017 – Apr 2017
Provides estimation for predicting the closure time, team assignment, service assignment for any incoming incident tickets based on the text content of the corresponding incident in real time.

Personalization for South Indian regional movie content (FastFilmz)   Sep 2017 – Dec 2017
Set up the data science team at Fastfilmz (Google Launchpad 2017). Built their in-house personalized recommendation platform. Prototype is available at https://github.com/tuhinsharma/imdb-rec-sys

Publications

1. Sahisnu Mazumder, **Tuhin Sharma**, Rahul Mitra, Nandita Sengupta and Jaya sil **(2012), "*Generation of Sufficient Cut Points to Discretize Network Traffic Data Sets*",** Proceedings of Third International Conference on ***Swarm, Evolutionary, and Memetic Computing*** **(SEMCCO)** published in LNCS 7677, pp. 528-539, by Springer-Verlag Berlin Heidelberg, 2012.
2. Rahul Mitra, Sahisnu Mazumder, **Tuhin Sharma**, Nandita Sengupta and Jaya sil **(2012), "*Dynamic Network Traffic Data Classification for Intrusion Detection using Genetic Algorithm*",** Proceedings of Third International Conference on ***Swarm, Evolutionary, and Memetic Computing*** **(SEMCCO)**, published in LNCS 7677, pp. 509-518, by Springer-Verlag Berlin Heidelberg, 2012.
3. **Tuhin Sharma,** Durga Toshniwal **(2014)**, **"** *A Generalized Relationship Mining Method for Social Media Text Data***",** Proceedings of **10th International Conference on Machine Learning and Data Mining (MLDM 2014)**, published in LNCS 8556, pp. 376-392, by Springer-Verlag Berlin Heidelberg, 2014.
4. **Tuhin Sharma,** Durga Toshniwal **(2016)**, **"** *A Generalized Framework for Quatifying Trust of Social Media Text Documents***",** Proceedings of **12th International Conference on Machine Learning and Data Mining (MLDM 2016),** published in LNCS 9729, pp. 692-713, by Springer-Verlag Berlin Heidelberg, 2016.

Patents

1. **Tuhin Sharma**, Krishna Kummamuru, Ravi Tejwani, Jeff Eisen, "**User experience using social and financial information**" (Patent Number - US20170236215A1)
    https://patents.google.com/patent/US20170236215A1
2. **Tuhin Sharma,** Bharath Ganesh**, "Use of asset and enterprise data to predict asset personality attributes"** (Patent Number - US20180060735A1) https://patents.google.com/patent/US20180060735A1
3. Dhandapani Shanmugan, Bharath Ganesh, **Tuhin Sharma,** Jothi Subramani**, "Cognitively inducing Sleep Cycles by leveraging Wearables"** https://patents.google.com/patent/US20190160251A1
4. Dhandapani Shanmugan, Bharath Ganesh, **Tuhin Sharma,** Gerin Jacob**, "Mitigation Of Perishable Product Waste Through Promotion Recommendation"**
    https://patents.google.com/patent/US20190156369A1
5. Dhandapani Shanmugan, Bharath Ganesh, **Tuhin Sharma**, Jothi Subramani, **"Enabling Wearables To Cognitively Alter Notifications And Improve Sleep Cycles"**
    https://patents.google.com/patent/US20190124032A1

Other Interests/Achievements

- *Datahack Summit Speaker 2019* – Presented a talk about "Federated Learning using Deep Learning" https://www.analyticsvidhya.com/datahack-summit-2019/schedule/hack-session-secure-and-privacy-preserving-deep-learning-using-federated-learning
- *O'reilly AI Conference Speaker 2019* – Presented a talk about "Anomaly detection in smart buildings using federated learning" https://conferences.oreilly.com/artificial-intelligence/ai-eu/public/schedule/detail/78152
- *ODSC Speaker 2018* – Presented a talk about "Hybrid Recommendation Systems in News Media using Probabilistic Graphical Models" https://confengine.com/odsc-india-2018/proposal/7098/hybrid-recommendation-systems-in-news-media-using-probabilistic-graphical-models
- *O'Reilly author* – Deep Learning for NLP with MXNet on AWS. Profile and articles can be found at https://www.oreilly.com/people/tuhin-sharma
- Won *First Runner Up at Hackasm 2017* (Ivanti), for developing a chatbot (using Deep Learning) to automatically detect problems raised by customers and raise corresponding incident tickets. Code is available at https://github.com/tuhinsharma/chatbot-incident-ticket
- Won *First Runner Up at Techathon 2015* (IBM India Pvt. Ltd.) for developing an app which creates automatic highlights of a cricket match using events mined from social media.
- *Excellence and Eminence Award 2015* (IBM India Pvt. Ltd.) **-** Critical and Timely contribution to content building on NGP platform.
- *Goyel Award 2014* **(**Indian Institute of Technology Roorkee**)** Topper of the batch of 2012-2014 in M.Tech (CSE**)** programme.