ASHWIN K. THILLAI NATARAJAN
A Data science practitioner with passion for analytics.

75, St.Alphonsus St., Boston, Massachusetts 01938   ph.-(617)-982-8863   email- thillainatarajan.a@husky.neu.edu

PROFESSIONAL SUMMARY

A former Statistics and quality planning engineer with newly acquired skills, an insatiable intellectual curiosity, and the ability to mine hidden gems located within large sets of structured, semi-structured and unstructured data. Able to leverage a heavy dose of computer programming and applied statistics with visualization and a healthy sense of exploration.

AREA OF EXPERTISE

Predictive and classification modelling, Text mining, Natural Language processing, Data cleaning and visualization, Machine learning, Data management, Warehousing, Relational database, Map reduce and distributed file system, Data analysis and business solutions, Statistical and risk modelling.

PROJECTS

Polymerase paper classification system- *New England Biolabs*   **spring 2015**

- Built a DNA-polymerases science-paper classification system from scratch using machine learning techniques implemented entirely using python 2.7.9 and python libraries such as numpy, matplotlib, psycopg2 and scipy.
- Applied text-mining principles and parsed large volumes of text data in to quantitative information. Extensive use of natural language tool kit python library.
- Implemented a python web crawler to access medical journals from National Library of medicine and PubMed repositories.
- Designed database to support data management of the paper-classifier system using PostgreSQL.
- Worked in integrating Ruby on Rails web application with paper-classification system.

Tree based prediction and sales forecast report- *Northeastern University*   **summer 2014**

- Developed a regression tree based prediction model using R programming for predicting sales of resale goods for a large retailer in Europe.
- Implemented bivariate variable reduction technique for feature selection.
- Built a time series forecasting system to predict future profit of the retail chain using spreadsheet apps.
- Suggested resale good with high sales potential and business prospects.

Analytical data management system for e-commerce web app- *Northeastern University*   **spring 2014**

- Developed a relational database system to support high-level analytic and reporting activities.
- Implemented PostgreSQL- SAS integration for carrying out analytical operations.
- Implemented database optimization techniques such as Indexing to improve performance.

SOFTWARE AND TOOLS

Python, R, SQL, SAS, GitHub, Linux command line, Ruby on Rails, Minitab, Microsoft Excel, Strata, SSIS, PostgreSQL, Weka Data mining, Python NLTK, Scikit-learn.

EXPERIENCE

Student Intern- *New England Biolabs, Ipswich, MA*   **January 2015 – June 2015**

- Developed personalized paper classification system for Biolab's proprietary web application.
- Report daily activities and findings to application and product development scientist on daily basis.
- Explore optimization techniques for improving the performance of polbase web app using ruby on rails application.
- Involve in developing data preparation and visualization systems for monitoring key process indicators.

Graduate student Assistant- *Northeastern University, Boston, MA*   **December2013- March 2014**

- Design and conduct statistical tests and models for different projects.
- Collaborate with team to collect data and generate reports.
- Conducted Preference ratio analysis to determine customer preference on two dependent library resource
- Used a custom-made data collection app integrated to internal database to build and manage data for various projects.

Graduate Engineering Trainee- *Mahle IPL Ltd, TN, India*   **August 2011- April 2013**

- Led Statistical Process Control team and reduced the overall machining loss by 10% by identifying contributing factors trough regression analysis and six-sigma.
- Oversaw Advanced Product Quality Planning (APQP) for clients including Toyota Corporation and Ford India Ltd.

EDUCATION

Masters of Science (IE) degree, Northeastern University, Boston, MA   **Expected August 2015**

Relevant courses: Engineering Probability and Statistics, Deterministic Operations research, Statistical method in Engineering, Database management system, Data Mining Techniques, Machine Learning

Bachelor of Engineering (Mechanical) degree, Anna University, India   **June 2011**

CERTIFICATION

HBX Credential of Readiness - Offered by Harvard Business School   **January 2015**
Financial Accounting, Business Analytics

Green Belt in Six-Sigma- Offered by AU-TVS institute of Quality   **June 2012**
 DMAIC principle, Design of Experiments, Statistical Modelling and Lean-Six sigma