Shivam Pahuja

Data Scientist

080-8887-1700   shivam.pahuja.ee@gmail.com
linkedin.com/in/shivampahuja

Career Profile:
Analytical and technology enthusiast with a keen eye for details, able to effectively coordinate tasks to accomplish projects with timeliness and creativity. Experienced in data science techniques of acquiring, cleaning, and building business solutions over various datasets.

Technical Skills:
Python (Flask, Sklearn, Pandas, NumPy, SciPy), Hive, Spark, MongoDB,
Elastic Search, MySQL, R
Tools: Tableau, Airflow, Kibana, Scrum Development, Docker, Git, Jenkins,
Pentaho (BA server, Data Integration Tool)

Professional Experience:

Data Scientist, Philip Morris International, Tokyo

March 2018 – Present
- **Price Recommendation Project:** Designed and developed a diagnostic tool which can capture possible interactions between various products using historical price and sales time series and the insights could be used in a actionable fashion (Tableau, Hive, Python, Docker)
- **Consumer Segmentation for Targeting**
- **Realtime Reputation Management**

Data Science Engineer, Rakuten, Tokyo

January 2017 – March 2018
- **Recommendation System**: Worked on merchant recommendation system using user behavioral data (R, Python, Spark, Word2Vec)
- **Look-alike Modelling:** Involved in Prediction Framework development by vectorizing user and item data sets to do user segmentation and create useful features which can be used for personalization and find look alike users to expand targeted marketing
- **Web Optimization tool:** Designed and developed API(python/flask) for web marketing tool capable of doing Targeting, A/B Testing with Bandit optimisation and providing user detailed reports of their experiment
- **Data Pipelining:** Created robust data pipelines for data transactions between different data sources using Airflow (Spark, Hive, Python, MySQL)

Data Science Engineer, Hikari Tsushin Inc., Tokyo

May 2015 – January 2017

- **BI Reporting Tool:** Integrated Pentaho Business Analytics Server with the existing internal sales server and created various analytical reports, dashboards and visualizations through it
- **Data Pipelining:** Extracted, manipulated, and merged large databases to create analytical datasets. (Python, Oracle, MongoDB)
- **Customer Ranking Model:** Created mathematical models using R (R studio) to uncover trends and relationships in data and design visual representations of data for clear and concise communication of findings. Continuously developed insights into customers and products to drive growth and communicated actionable learnings to drive positive change.
- Worked on internal employee survey analysis and Hourly Employee Performance reporting system (QlikView, VBA)

- Automated Financial reports. (QlikView, VBA)

Trainee, Hikari Tsushin Inc., Tokyo

Nov. 2014 –May 2015

- Led a team of 7 to build and maintain a project management system using SalesForce platform
- Handled the whole backend design and development

Web Solution Developer ,Times Internet Limited

June 2013 – November 2013 (Internship)

- Analysed, designed and developed a complex User Profiling system which can handle large user data (used JAVA, Mongo DB, Spring Framework)
- Developed a scalable Newsletter mailing service which can track user interaction data (used JAVA, Spring Framework)

Computer Vision Developer , Hi-Tech Robotics Systemz Ltd.

Dec. 2012 – March 2013 (Internship)

- Build a shadow removal algorithm which can enhance surveillance systems and consumer images (C++, OpenCV)

Other Projects:

- Hand gesture recognition application
- Multiple Data Compression and encryption using iterative techniques (C/C++)
- Linux System Administration and networking
- Android Application Controlled Bluetooth enabled Bot

Certifications:

- Certified Scrum Developer (http://www.scrum.org/user/348965)
- Kaggle R Tutorial on Machine Learning [DataCamp, License 1832ff8571a3536f6695dbc70c014396c326cb9f]
- Oracle Certified Associate, Java SE 7 Programmer [Oracle, License OC1487578, May 2015]
- Oracle Database 12c: SQL Fundamentals [Oracle, License OC1487578, August 2015]
- GAIT Certified IT Professional – Silver [日本サード・パーティ(株), License D0400109mp0000000053, Sep. 2014]
- Data Analysis and Statistical Inference [DataCamp, License fd44d08e475e5f22ecd90a8e12f4b94d96be1053]

Education:

Delhi College of Engineering (Delhi Technological University), India
Bachelor of Engineering Electronics and Communication 2010 – 2014