

Sumeet Bhadana

<https://summii.github.io>

Email : sumeetbhadana99@gmail.com

Mobile : +91 9739026151

EDUCATION

- **Dr. M.G.R. Educational and Research Institute**

Chennai

Bachelor's Degree in Information Technology; CGPA: 7.31

Aug. 2009 – Dec. 2013

EXPERIENCE

- **Semantics3**

Bengaluru

Engineer

Dec 2013 - Present

- **Crawl:** Develop highly scalable and efficient web crawlers in perl, as well as worked on content extracting and validation techniques for ecommerce data
- **Url Canonicalization:** Develop url canonicalization techniques for e-commerce sites; primarily used for Url clean-up.
- **Data Mining:** Create data mining techniques to build scalable and robust data extraction platforms using Regular Expressions.
- **Custom Crawl/Custom Ops:** On demand perl script to extract information from ajax/javascript enabled pages. Generated reports on E-commerce sites based on product data i.e variations, javascript/ajax enabled, price etc
- **Wrapper:** Writing wrappers using regular expressions for crawl/discovery of daily products price refresh.
- **WrapperQA:** Generating reports on sites data in our db to check data validation errors.
- **Pricing Suggestion Engine:** Build an algorithm that automatically suggests the right product prices given descriptions of products, including details like product category name, brand name, and item condition

- **Pace Wisdom**

Bengaluru

Web Developer Intern

Aug 2013 - Nov 2013

- **Web development:** Design, build, or maintain web sites, using authoring or scripting language.
- **Design:** Create web models or prototypes that include physical, interface, logical, or data models.

PROJECTS

- **Recommendation System and EDA:** Developed/Explored a basic recommender system with Python and pandas.
- **Real Time Product Data Scrapper:** Build, Develop and Deploy a real time product data scrapper for e-commerce in flask.
- **Semantic Aware Metadata Organization Paradigm in Next-Generation File System :** This project proposes a novel decentralized semantic aware metadata organization which explores semantics of files metadata to judiciously aggregate correlated files into semantic aware groups by using information retrieval tools to limit the search scope of a complex metadata query.

PROGRAMMING SKILLS

- **Languages:** Python, Perl, Numpy, Pandas, Regex, C
- **Frameworks:** Scikit, Pandas, Numpy, Git, Flask

Database Systems: MySQL
Technologies: AWS, Vim, Tmux

EXTRACURRICULAR

- **Sports:** Basketball, Football
- **Others:** Worked with Magic Bus(NGO) in rural areas on various programs.