

# Sumeet Bhadana

<https://summii.github.io>

Email : [sumeetbhadana99@gmail.com](mailto: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

- *Data 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.

- **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

---

- **Slack bot:** Build a slack that return top Reddit story from any subreddit.
- **Car Price Prediction Model:** Build, Develop and Deploy a Machine Learning Model to predict cars price using Gradient Boosting.
- **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, SQL, Perl, Numpy, Pandas, Regex

**Technologies:** AWS, Vim, Tmux

## EXTRACURRICULAR

---

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