

**Sachin Kalsi**, passionate programmer

[sachinkalsi15@gmail.com](mailto:sachinkalsi15@gmail.com),

+91 9945746470

[Portfolio](#) | [Blog](#) | [LinkedIn](#) | [Github](#) | [Stackoverflow](#)

---

Senior Data Scientist with 4.5 years of experience in Software development & 2+ years of experience in Machine Learning focused primarily on NLP

## EXPERIENCE

### **DRAUP (draup.com)**

*Senior Data Scientist*

- Built NLP based search engine including custom entity recognition.
- Worked on generic scrapers, browser extensions, Information retrieval from unstructured data, pattern identification.

### **Calient Technologies**

*Software Engineer*

- Analyze network device logs & classify the issues.
- Algorithm design to auto discover optical switches port.

### **Cogoport**

*Software Developer*

- Crawl websites, data analysis, DB & REST API development.
- Mentoring Interns, Data interpretation for sales team.

### **Techniche e-commerce**

*Full Stack Developer*

- Built e-commerce websites, full stack development.
- Train & mentor freshers.

### **Hewlett-Packard (HPE)**

*Technical Consultant-II*

- Built internal Java applications

## EDUCATION

**Bachelor of Engineering (B.E.) in Computer Science (2009 - 2013)**

*Visvesvaraya Technological University (VTU)*

## SKILLS

**Programming languages:** Python, Java(Fluent, Prior Experience), Javascript (Proficient)

**Libraries:** Pandas, Scikit-Learn, NLTK, Gensim, Seaborn, tf.keras v2.2, PySpark (Basic)

**Database:** MongoDB (Fluent, Prior Experience), MySQL (Basic)

**Machine Learning:** KNN, Logistic regression, Random Forest, SVM, Ensemble models, Unsupervised algorithms, Word2Vec, Basic Stats

**Deep Learning:** LSTMs, Seq2Seq

**Other tools & techniques used:** [DBPedia](#), Databricks, Rasa.ai, BERT (base)

## PROJECTS

### ❖ Natural Language Search Engine

- Redirecting the user queries (in natural language) to one of the relevant pages along with the identification of entities such as job title, skill etc
- Various models have been trained, developed, tested and deployed for intent classification & named entity recognition (NER).

### ❖ Cluster Similar News at Scale

- Summarise the news description and encode into high dimensional meaningful vectors.
- Built custom ML models to cluster the similar news of a given company by taking the window of the last 8 days' news and to pick the relevant news from each of the clusters.

### ❖ Generic web scrapers

- Built a generic web scraper to extract various fields from a given university URL which reduced the human efforts by 60%
- Built chrome extensions which helped in the annotation of various HTML elements which resulted in more accurate data extraction

### ❖ Business function classification

- Classify the outsourced projects of a company into various categories
- Achieved a precision of 0.85 with well calibrated probabilities.

### ❖ Error Log Analysis

- Analyse the error log files produced by network devices
- Classify and assign the priority for the logs