

# Somin Wadhwa

---

## CONTACT INFORMATION

E-mail: [sominwadhwa@gmail.com](mailto:sominwadhwa@gmail.com)Webpage: [sominwadhwa.github.io](https://sominwadhwa.github.io)GitHub/Kaggle: [/sominwadhwa](https://github.com/sominwadhwa)LinkedIN: [/in/sominwadhwa](https://in/sominwadhwa)

## EDUCATION

### Bachelor of Technology in Computer Science & Engineering

Maharaja Agrasen Institute of Technology

Guru Gobind Singh Indraprastha University, Delhi, India

2014 – 2018

CPI: **79.2%**

## RECENT EXPERIENCE

### • Indraprastha Institute of Information Technology, Delhi (IIIT-D)

*Research Intern*

July, 2018 - Present

Complex Systems Lab, Center for Computational Biology

June 2017 - March 2018

**Advisor:** Prof. Ganesh Bagler

– **Current Work:** My current work focuses on creation of *“BitterSweet: A resource to explore and predict taste information in small molecules”*. (<http://cosylab.iiitd.edu.in/bittersweet>)

– **Previous Work:** Pursued an internship where I worked on devising new methods to predict phenotypic side effects of drugs using existing data (SIDER4 & DrugBank). The problem was formulated as an extreme multiclass-multilabel classification problem with severe class imbalance in the datasets. The work conducted was culminated in the form of a research article (published) and the code was documented & is now open-sourced on Github ([code](#)).

### • All India Council for Technical Education (AICTE), Govt. of India

*Research & Development Intern*

October 2017 - March 2018

**Advisor:** Dr. N.H. Siddalingaswamy (Director, AICTE)

– **Work:** Lead a team of 5 with a project budget of \$4600 (Rs. 300,000) sponsored by the Ministry of Human Resources Development, Government of India. We curated and analysed graduate employment statistics datasets of several years and developed dynamic analytic dashboards to aid AICTE in granting approvals to higher education institutions. ([code](#))

## PUBLICATIONS

- Tuwani R, **Wadhwa S**, Bagler G (2018) BitterSweet: Building machine learning models for predicting the bitter and sweet taste of small molecules. bioRxiv 426692 (*preprint*)  
doi: <https://doi.org/10.1101/426692>
- **Wadhwa S**, Gupta A, Dokania S, Kanji R, Bagler G (2018) A hierarchical anatomical classification schema for prediction of phenotypic side effects. PLOS ONE 13(3): e0193959  
doi: <https://doi.org/10.1371/journal.pone.0193959>

## SELECTED PROJECTS

All of my projects (including the following) are available on [github.com/sominwadhwa](https://github.com/sominwadhwa)

### • Visual Question Answering through Modal Dialogue:

A semester long B.Tech project based on the application of *Malinowski et al.* (CNN + LSTM based models) on v2 of the VQA dataset released in April 2017. I have also documented and made the entire process reproducible with a single click in the form of a [featured blog post](#). ([code](#))

### • Kaggle Repository:

An ongoing (2+ years) collection of kernels (implemented using IPython notebooks) designed using datasets obtained from Kaggle for practise as well as competitions. Projects in this repository demonstrate the implementation of several algorithms and data visualization techniques. ([github-repo](#), [kaggle profile](#))

### • TheTwitterPolice:

A basic analysis & visualization of the Indian law enforcement activity on Twitter. Collected data for different cities (BeautifulSoup & Selenium), stored them in a database (MongoDB), analysed (sentiment analysis, basic statistics etc) & displayed the results graphically through a flask web-application. ([code](#))

### • Image Apportionor:

A simple clustering based image segmentation project. Implemented the k-means clustering algorithm, from scratch, in Python as part of the academic summer activity requirement in my sophomore year. ([code](#))

## SKILLS

- **Languages:** Python (proficient), Java (familiar), bash scripting, SQL, L<sup>A</sup>T<sub>E</sub>X.
- **Frameworks & Libraries:** Tensorflow, keras, PyTorch, scikit-learn, SpaCy, NLTK, Matplotlib, Plotly, MongoDB, Flask.
- **Relevant Classes Taken:** Data Structures, Algorithms Design & Analysis, Probability & Curve Fitting (Applied Mathematics-IV), Advanced DBMS, Machine Learning, Soft Computing.
- **MOOCs:** Deep Learning Specialization by Andrew Ng (5 courses, [certificate](#)), Algorithms Design & Analysis by Tim Roughgarden ([certificate](#)), Machine Learning by Andrew Ng ([certificate](#)).

## ACHIEVEMENTS & OTHER ACTIVITIES

- **Outstanding Achievement Award:** Conferred by the Computer Science & Engineering department at MAIT, Delhi for my contributions to the institute.
- **Best B.Tech Project:** My senior year project on Visual Question Answering was selected among the top four out of the 60+ projects reviewed by the department.
- **Smart India Hackathon 2017, MHRD, Govt. of India:** Lead a team of 6-members advised by Dr. Sambuddha Roy (Principal Data Scientist at Microsoft, Seattle) and won first prize with a total reward of \$3000 (Rs. 200,000) awarded by Government of India and MAIT, Delhi. As part of the winning team, we were also provided with an opportunity to take our prototype forward and this resulted in a 6-month long internship with the All India Council for Technical Education (AICTE).
- **Secretary, Association of Computing Machinery:** Served in the capacity of Secretary of 80+ strong team of ACM-MAIT Student Chapter during 2015-2016.
- **Blogging:** Maintain an active blog at [sominwadhwa.github.io/blog](https://sominwadhwa.github.io/blog) to document some of my experiences & selected projects (for reproducibility).