

# Shubhanshu Gupta

Email : shubhanshu.gupta93@gmail.com

Mobile : +65-82923526

[shubhanshugupta.com](http://shubhanshugupta.com)

[LinkedIn](#) | [Medium](#) | [Github](#) | [Google Scholar](#) | [Stack Overflow](#)

## EDUCATION

---

**National University of Singapore (NUS) | NUS Computing and Business School** Singapore  
Master of Science in Business Analytics; GPA: 3.71/5.0 Jul'18 - Sep'19

**Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT)** Gandhinagar, India  
Bachelor of Technology in Information and Communication Technology; GPA: Jul'11 - May'15  
7.58/10.0

## PATENT & PUBLICATIONS

---

Shubhanshu Gupta et al. **Machine Learning Modeling to Identify Sensitive Data, U.S.A & Singapore, Patent Pending** USA, Singapore  
Sep'21

**Music Data Analysis: A State-of-the-art Survey** Harvard University,  
Academy of Science and Engineering (ASE), Fourth International Conference on USA  
BigData (Acceptance Rate ~8.5%) Dec'14

**Social Data Analysis: A Study on Friend Rating Influence** VmWare, India  
WWC Connect Conference Mar'17

## EXPERIENCE

---

**Citibank, Global Consumer Banking (GCB) | Data Scientist** Jul'20 - Present  
(Singapore)

Driving the initiatives under **AI/ML Center of Excellence (CoE) as the Technical Lead**, by collaborating with cross-functional teams & stakeholders. This involves understanding & crafting use cases, building POCs through rapid prototyping in areas such as Data Governance, Wealth Management, Infrastructure Monitoring, & Customer Experience.

### PII Detection Framework (Pending Patent) under Data Governance

- Designed & Implemented ML based PII Detection Framework. It prevents sensitive data exposure in data-warehouse, using an amalgam of **probabilistic classification models** on column metadata and **Natural Language Processing (NLP)** on column content.
- The framework identifies misclassified columns in the data warehouse and assigns PII sensitivity to the new columns. In addition to being scaled to all Citi markets globally, the framework is enabling data governance teams to improve their capacity & efficiency by **30x per market**.
- The project won the first prize (across 70 teams) in a company-wide competition in APAC, for solving one of the most critical pain points.

Established a **Machine Learning Kickstarter** program to **mentor** and guide Citi colleagues, learn Machine Learning.

**Ad Tech | Marketing Data Scientist** Dec'18 - Jul'20  
(Digitas - Publicis Groupe, Dentsu Aegis Network) (Singapore)

(Digitas - Publicis Groupe) Enabled clients in the FMCG & Hospitality sectors to understand their audiences in new markets. **Increased Engagement rate** of campaigns by 30% and **decreased media cost** by 21%. Used

**DeepLearning** library for Named Entity Recognition, and **NLP** to extract highly collocated words, to build audience insights.

(**Dentsu Aegis Network**) Strategized winning the right ad impression at the right bid price, using Machine Learning in Programmatic Advertising. Built **decision tree models** to identify & target the audience segment based on their behavior, intent & interests. Used **A/B tests** to evaluate the performance of campaigns.

## **Early & Mid Stage Growth Startups | Software Development Engineer**

Jun'15 - May'18  
(India)

Handled various engineering roles and projects including product backend, devOps, and analytics. Agile methodology, clear communication with stakeholders, and continuous development in short sprints, are some of the values I accrued while working with small and dynamic teams in growth startups.

- (**Truebil**) Designed and implemented Truebil's Progressive Web App (PWA) backend in Python Django which **increased the overall Buyer Lead Conversions by more than 2x**.
- (**Truebil**) Integrated CRM (detailed blog post on <https://bit.ly/shubhanshu-crm-blog>) which brought departments working in silos such as post-sales, customer-service, etc onto one platform, **leading to increased customer satisfaction**.
- (**LaughGuru**) Designed and implemented the user and sales analytics reporting platform. The former helped the customers visualize their progress and the latter, helped the management **visualize the sales KPIs**.
- (**MuSigma**) Restructured the rewards and loyalty scheme for a telecom client, **leading to personalized rewards to customers**.

## **TheCollegeStore - Bookly Web LLP | Founder**

Oct'13 - May'15  
(India)

- Founded an online marketplace for students to buy and sell books. **Received seed funding and incubation**.
- **Scaled up to 35 colleges with 1000 paid MAUs** in less than six months. Mentored and led a 15 member team for product and business development.

## **SKILLS**

- **Machine Learning:** Natural Language Processing, Time Series Forecasting, Recommendation Models, Marketing Analytics, Attribution Modeling, Regression, Classification, Clustering, Data Mining, Cleaning & Visualization (Matplotlib, Plotly, Tableau)
- **Databases & Big Data Technologies:** MySQL, postgresSQL, couchDB, Google BigQuery
- **Programming:** Python (Pandas, NumPy, scikit-learn, Boto, Spacy, NLTK), SQL, Django

## **EXTRA-CURRICULAR**

### **Business Analytics Club | Founding President**

Founded Business Analytics Club at NUS, conducted data jamming sessions, career guidance activities, and alumni chat sessions.

Aug'19 - Jul'20  
(Singapore)

### **Hong Kong Smart City Datathon | Global Top 20 Finalists**

Designed Optimal Turn Around Time solution for Hong Kong Airport Authority using 3 months of airport data.

Oct'18  
(Hong Kong)