

# Karan Arya

- Pune, India • [karan.arya4196@gmail.com](mailto:karan.arya4196@gmail.com)
- +91-7390082180 • [LinkedIn](#) • [GitHub](#) • [Portfolio Website](#)

## EDUCATION

- Indian Institute of Technology (Banaras Hindu University), Varanasi**, B.Tech + M.Tech 2013-2018
- GPA: 9.32/10. Rank: 2. Major: Pharmaceutical Engineering. 5-year Integrated Dual Degree (B.Tech + M.Tech).
- Symbiosis Institute Pune**, PGDM, Data Science and Statistics 2019-2020
- GPA: NA/10, Rank: NA.

## EXPERIENCE

- Wolters Kluwer**, Data Scientist I Dec 2018-Present
- Worked on NLP and machine learning modeling to develop an AI-powered product - Bill Analyzer.
  - Developed a full-fledged Apache Spark version of the Bill Analyzer model to manage big data.
  - Involved in the predictive modeling of a lawfirm recommendation system for clients using deep learning and NLP.
  - Domain specific keyword extraction using custom algorithms from WordNet and spaCy.
  - Developed a micro-model using Bag of Words Classifier for binary classification of flagged invoices based on bigrams and trigrams.
  - Built a Conditional Random Field classifier micro-module for identifying tasks specific to a category.
- Purdue University**, Research Assistant Aug 2018-Nov 2018
- Joined as a PhD candidate focusing on the applications of machine learning in Chemistry but dropped out due to father's critical health condition.
- University of Cambridge**, Informatics Intern May 2017-July 2017
- Modeling and target prediction of Afrocancer compounds.
  - Cancer classification of cellular data using machine learning.
- University of Würzburg**, DAAD (German Academic Exchange) Intern May 2016-July 2016
- Optimization of iClick reactions using analytics and modeling.
- Treebo Hotels**, Business Analytics Intern May 2015-July 2015
- Contributed to the business development and sales inventory analytics.

## SKILLS

- **Business Analytics & Data Science:** Excel, Tableau Basics, NLP, Deep Learning, Machine Learning, Statistics.
- **Libraries:** Sklearn, Pandas, NumPy, TensorFlow, Keras, NLTK, spaCy, Spark-Mllib, Spark-NLP, WordNet.
- **Languages:** Python, Spark, SQL Basics, HTML5, CSS3.
- **Soft Skills:** Problem Solving, Teamwork, Communication, Accountability, Leadership, Presentation.

## COURSES AND CERTIFICATIONS

Basics of Statistics	Python Programming	Neural Networks and Deep Learning
Introduction to Data Science	Machine Learning with R and Python	CNNs and Sequence Models (RNNs)
Data Structures and Algorithms	Advanced Statistics and Data Analysis	TensorFlow Specialization
Introduction to R Programming	Emerging Trends in Data Science	+ 3 more Coursera certifications

## AWARDS

- Recipient of the Ross Fellowship at Purdue University, USA and DAAD WISE Scholarship, Germany.
- Awarded US \$5000 for Bachelors Thesis by Govt. of India. Secured the IIT Merit-cum-Means Scholarship from the Govt. of India, HDFC Bank Scholarship, BHEL FAEA and GATE Fellowship.
- Held the position of Marketing Head and Scientific Coordinator at the annual festival of IIT BHU Varanasi.
- Led the Spirit '16 blood donation campaign in Uttar Pradesh.

## COMPETITIONS AND PROJECTS

- Developed a responsive front-end demo website using HTML5 and CSS3 (Repository on GitHub).
- Drug repurposing to target anti-tubercular enzymes utilizing drug discovery algorithms and network analysis.
- As a Knowledge Associate, managed a team of 30 interns in order to ensure the procurement of quality educational content.
- GRE Score: 328/340. TOEFL Score: 112/120.