Research Engineer - Data Science(NLP)
Product Labs, IIIT Hyderabad

ABHISHEK ARORA

+91 9711408264
abhisheknsit15@gmail.com

ACADEMIC QUALIFICATIONS

Degree	Year	Institute	Board/University	%/CGPA
B.E	2015-19	Netaji Subhas Institute of Technology	University of Delhi	8.61/10
Class XII	2015	Delhi Public School Rohini	CBSE	94.6%
Class X	2013	Delhi Public School Rohini	CBSE	9.8/10

- Among **top 10%** students in the department at NSIT.
- Institute **Gold Medal** for **Academic Excellence** awarded by DPS Rohini for being a scholar badge for **6 consecutive years**.

ACHIEVEMENTS

- **Global Winner of the Google Code In 2011 (Youngest Winner)**
  - Google Code-In 2011 was an international competition organised by Google for pre-university students. The competition saw a participation from 56 countries. I worked with 5 Open Source Organisations and was **ranked #2** (youngest of all winners), received **$500** and a fully sponsored trip to **Google's HQ (Mountain View , California).**
- **Research Publications**
  - "Applications of Big data in Real World" in **IEEE Xplore** under Prof. Vidhi Khanduja. **2017**
  - Developed an optimisation algorithm, "A Novel Extremum Partitioning Duty cycle based P&O Technique for Efficient MPPT of Photo-Voltaic Systems" under Prof. K.P.S Rana. (Currently under review) **2019**
- **International Student Travel Award** for **KDD 2017**
  Top Data Science conference held at Canada in August 2017
  - Received **$1000 scholarship** and an opportunity to travel to **Canada** to meet, interact with leaders and professionals in the data science community, attend various workshops and the emerging research in this field.
- Finalist in Building Better Villages Case study organised by **Ministry of Panchayati Raj, (Govt. of India)**at IIITD.

TECHNICAL SKILLS

- Passed **NPTEL** certification on **Programming , Data Structures and Algorithms.**
- **Programming Languages**: C/C++, Python, Java(Beginner), LateX, HTML/CSS.
- **Software Packages & Libraries**: MATLAB ,Django, Doxygen, Libraries - Numpy, Pandas , Quandl, Scikit Learn, Scipy , Git..

WORK EXPERIENCE

- **Research Engineer - Data Scientist** at **Product Labs, IIIT Hyderabad (January 2020 - Present)**
  Product Labs aims to build commercial products by the application of research in the domain of CV and NLP.
  - Working on a deep learning based document question answering algorithm which can extract answers from a large corpus of documents to assist users with their natural language queries and increase the efficiency of scanning hundreds of pages in various domains like legal, banking, finance, healthcare etc.
- **SDE - Data Engineering** at **HT Digital Streams (July 2019 - January 2020)**
  - Automated integration and collation of various news source API's to dump real time data into Redshift database which was utilised for analytics.
  - Completed a POC using NLP to classify live news feed data (collected from news sources at various platforms like twitter, ANI direct feed, Reuters etc.) as breaking news and trigger alerts for editors which was later deployed on hindustantimes.com for readers as well.
  - Data anomaly detection in the data pipeline to improve data driven decision making.

INTERNSHIPS AND POSITION OF RESPONSIBILITY

- **Summer Intern at Telecommunications Consultants India Limited, 2018.**
  *Implemented a **Smart Water Management system** using IoT and AI (proj. under Government of Telangana)*
  - Trained a **DBN neural net** to predict total pump energy rate for each channel for equitable distribution of water.
  - Implemented a **graph based network** for all IoT sensors to calculate minimum cost path, detect leakages, pressure, flow analysis. Used graph algorithms to recover the network in case of water leakages , pipe burst etc.
- **Research Intern at Indian Institute of Technology (IIT) Delhi, 2017**
  *Live project: **Increasing TB diagnosis rate** in India using **predictive modelling***
  - Working to build a predictive model to predict the spread of Tuberculosis in India and risk in each district of India using factors like population, income, re-emergence case rate, mortality rate etc. using sklearn and pandas
  - Aim to increase the diagnosis rate (currently ~55%) so as to eradicate TB by 2025.
- **Winter School on 'Digital and the Everyday: from codes to clouds' at IIIT B, 2017**
  - Selected among 30 participants from India to discuss and understand the socio algorithmic processes, data protection and privacy, digital finance, machine zone, e-Health and **Ethics in AI** and their impact on millions of users.
  - Created a technology enablement plan for Kathputli Colony which artists recognized at international level.
- **Secretary, Academic Affairs, Student's Council of NSIT**
  - **Lead** the Student **Mentorship** program for **650 freshers** with over **200 mentors** to ensure their smooth transition to college life and they can make informed decisions.

- **Account Administrato**r - Sahana Software Foundation (FOSS organisation).

PROJECTS

- Built a **blog application** from scratch using Express, MongoDB, and Semantic UI , Node JS and RESTful routing
- **Voice & Sentiment classifier -** Trained an ML model based on a dataset consisting of 3168 recorded calls to predict the gender, sentiment and extract key properties using acoustic properties of voice and Speech. The model achieved an accuracy of 89%.
- **Help Desk Chat Bot -** Built a NLP based chatbot for automating help desk scenarios.
- **Project** on "Arduino Nano - **Mini Voyager** Project" at **CEDT , NSIT.**
  - Understood different concepts of programming with Arduino Nano and implemented a number of techniques like charlie multiplexing, debouncing, seven segment display.
  - Designed a 'Catch the LED' game using switches , potentiometer and a seven segment display to display the result.