

VISHAL BHATIA

Phone: 9013464101 | Email: vishal.bhatia789@gmail.com

Address: 6-B Gopal Park, 1st Floor, Near Ram Nagar, Krishna Nagar, East Delhi, 110051, India

CAREER OBJECTIVE

To make a major contribution in the field of technology and succeed in an environment of growth & excellence. Earn a job, which provides me job satisfaction and self-development and helps me achieve personal as well as organizational goals.

EDUCATIONAL QUALIFICATIONS

Degree	Discipline	Institute	Year	CGPA / Percentage
Post-Graduation Diploma	Data Science/ Big Data Analytics	Centre for Development of Advanced Computing (C-DAC), Noida	2020	78.38%
Bachelors of Technology (B.Tech)	Electronics & Communication	Jaypee Institute of Information Technology, Sector 62, Noida	2017	7.1 /10
Senior Secondary	CBSE	S.R.D.A.V Public School, Delhi	2013	80.00 %
High School	CBSE	S.R.D.A.V Public School, Delhi	2011	9.2/10

WORK EXPERIENCE /ACADEMIC TRAINING

Innodata Inc., Noida, UP

[March '20 – Present]

Designation – Software Engineer (Data Science)



As a Data Science Specialist, I am responsible for **automating** the workflow of various in-house processes of the company. The company majorly serves in the **publishing domain** and has lots of data to work on. Some of the projects that I have worked/ am working on are:

1. **iXBRL Automation Tool** – Designed a tool that uses **FastText** Machine Learning model (developed by Facebook), to train and predict the iXBRL tags for the input files in different formats like docx, pdf, rtf and txt. ML libraries like **NLTK**, and **Scikit-Learn** were used. Achieved an **accuracy of around 82%** by tuning various hyperparameters like 'epoch', 'learning rate': 'wordNgrams', 'bucket', 'dim', and 'loss'. Checked accuracy using **sklearn metrics** like **classification_report**, **accuracy**, **precision**, **recall scores** and also verified by manual testing. The tool saved a lot of manual efforts and time of the company.
2. **Primo Client Software** – Responsible for **UI/UX development** of Primo Client application used for handling various **data conversion** processes. Here, I got a chance to upgrade my core **Python programming** skills by putting various python libraries like **Tkinter**, **sys**, **requests**, **os**, **shutil**, **glob**, **json**, **datetime**, **random**, **PyQt5**, **PyQt5-tools**, **xml.etree.ElementTree**, **urllib.request**, **Pandas**, **Numpy**, **configparser** and **watchdog** into practice.
3. **Preprint Forward Flow Tools** – Developed preprint tools in **Python** which facilitates Forward Flow for Medical, Chemical, Biological and General journal/ research papers and archives. Worked on **Flask**, **Django REST API** and **Postman** tool for **managing the API** requests. Learned to handle file formats like **XML**, **HTML**, and **JSON**.

C-DAC, Noida (Academic Training)

[Sep'19 – Feb'20]

Completed a 6-months Post Graduation Diploma course in Big Data Analytics from C-DAC Noida. Have undergone professional training on **Big Data Technologies**, **AI/ML**, **Python**, **SQL**, **MongoDB**, **R**, **Java**, **Linux**, **Tableau**, **Advanced MS Excel** and **cloud computing**.



Phronesis Partners Ltd., Noida, UP

[Sep '18 – Sep'19]

Designation - Research Associate

It is a multi-national market research company headquartered at Singapore. I was a part of Competitive Intelligence team, which majorly handles **data research**. As an Analyst, I did in-depth interviews with certain C-level executives of various MNCs, performed **exploratory data analysis** (EDA) of client specific data for **market forecasting** using **MS excel**, and did Data Visualization using **Tableau**.



P&S Market Research Pvt. Ltd., Noida, UP

[May '17-Mar '18]

Designation - Research Analyst

It was a client facing role where I performed **primary and secondary research** talking with many Industry Leaders. Some of the key skills I developed are as follows: -

1. Market Engineering, Forecasting & Consulting using **MS Excel Data Analytics**
2. Preparing professional business research reports using **MS Word and PowerPoint**
3. **In-Depth Analysis** of various industries (such as ICT, Telecom, Semiconductors, Finance, Banking, FMCG, etc.)



SKILLS AND ABILITIES

- **Frameworks & Technologies:** Hadoop, Spark, Kafka, Tableau, Spotfire, PyCharm, Jupyter Notebook, Spyder, MS Word, Excel, and PowerPoint.
- **Python Libraries:** NumPy, Pandas, Scikit-Learn, TkInter, Matplotlib, Gensim, TensorFlow, Keras, SpaCy, XML ElementTree, urllib requests, configparser, PyQt5, NLTK, pytesseract, python-docx, sys, os, shutil, glob, json, datetime, and watchdog.
- **Other Known OS, Tools and Services:** Linux, MATLAB, Flask REST API, Postman and AWS Cloud Computing (RDS, S3 and EC2)
- **Programming languages:** Java, Python, R and SQL
- **Databases:** Oracle Database, MS SQL server, MySQL and MongoDB
- **File Formats worked with:** XML, JSON, PDF, HTML, CSV, ePub, xls, xlsx, docx, rtf, and txt
- **Hard Skills:** Coding, Data Analysis, Database Knowledge and Data Visualization.
- **Soft Skills:** Time Management, Teamwork, Good Communication and Sound Reasoning

KEY PROJECTS UNDERTAKEN

Projects in C-DAC:

1. Exploratory Data Analysis (EDA) and Text Analysis of Zomato Restaurants in India

The aim was to conduct descriptive, predictive, and prescriptive analysis of diverse Zomato Restaurant data set. It required a model that could help entrepreneurs in making sound business decisions, helping them decide what cuisine to choose, what kind of ambiance do people like, and what are the expectations of a customer from a new restaurant. We planned and executed our workflow in following way:

Live Data Extraction through Zomato API → Preprocessing → Model Fitting and Predictions → Validation and Scoring → Text Mining (tokenizer, bi-grams, trigrams, etc.) → Data visualization using Python → Data visualization using Tableau

2. C-DAC's Chatbot Assistant

Designed a chatbot for assisting C-DAC aspirants by answering all FAQ's that usually arise before taking admissions. This project is based on Natural Language Processing, and uses techniques such as LSTM, Lemmatization and Stemming.

Bachelors level (B.Tech) projects:

Automatic as Well As Manual Heating & Cooling Jacket

- The aim of the project was to invent a system, which cools and heats the body as per the requirement. Several units of setup including relays, temperature sensors, Bluetooth and UV sensors were embedded in a jacket. The command received from mobile Bluetooth was used to set the temperature required by the user. The system was also designed to monitor temperature from sensors and adjust the temperature automatically. The Chief Guest of our college exhibition, which was named **Cyber Srishthi** appreciated this Project. Presented in **IEEE conference titled INDIACom2017** held on March '17, the research was accepted for publication. The value added by this project is that it saved a considerable amount of electricity and money and is emission free (or eco-friendly).

Research and Publication

Paper titled "**TEMPERATURE CONTROL AND SOLAR RECHARGEABLE WEARABLE WITH ACCESSIBILITY USING NICHROME WIRE**" was presented in **INDIACom conference held in Mar'17** and is accepted for being published by **IEEE**.

Proceedings of the 11thINDIACom; INDIACom-2017; IEEE Conference ID: 40353 2017 4th International Conference on "Computing for Sustainable Global Development", 01st - 03rd March, 2017 Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi (INDIA)

Copy Right © INDIACom-2017; ISSN 0973-7529; ISBN 978-93-80544-24-3

ACHIEVEMENTS & EXTRA CURRICULAR

- Scored **highest marks** (in the 1st semester) in PG-DBDA course in C-DAC.
- Ranked **Top 25%** in TestDome's **MYSQL AND SQL challenge** - <https://app.testdome.com/cert/acff5c37769f465b8ea0def7d53e7149>
- Cleared **HackerRank's Python (Basic) Challenge** - <https://www.hackerrank.com/certificates/2a09fbafb2a6>
- Cleared **HackerRank's SQL (Basic) Challenge** - <https://www.hackerrank.com/certificates/f778fc53555d>
- Completed **Intro to Machine Learning** course at **Kaggle**
- Completed **Intermediate Machine Learning** course at **Kaggle**
- Have undergone **2 years of formal training** in Classical Music and Harmonium.



<https://www.linkedin.com/in/vishal-bhatia01>



<https://www.kaggle.com/vishalbhatia01>



<https://github.com/vishalbhatia1>