

Address

545 KA/SV-019, Sagar
Vihar Colony, Para
Road, Rajajipuram,
Lucknow-226017,
Uttar Pradesh, India.

Linkedin ID

www.linkedin.com/in
/gauravdubey-0703

Cell Phone No.

+91 700 7972332
+91-8764342247

Mail ID

gauravdube07@
gmail.com
imgauravdubey@
gmail.com

Programming Skills

Python
Java
R- Language
Machine Learning
AWS (Amazon Web
Services)
Hadoop
Apache Pig
Apache Sqoop
Apache Spark
Apache Hive
Scala
MS-SQL-Server(2012)
Microsoft office
Latex

Area of Interest

Data Science
Big Data Analytics
Machine Learning
Deep Learning
IoT
Auctorial Science

OS Preference

Windows ★★★★★
GNU/Linux ★★★★★

GauravDubey

Data Scientist, Big Data Analyst

Education

- 2017 - 2019 **MSc. in Computer Science(Big Data Analytics)** [Central University of Rajasthan, Ajmer, Rajasthan, India](#)
Specialization: Big Data Analytics.
- 2012 - 2015 **Bachelor of Science** [University of Lucknow, Lucknow, Uttar Pradesh, India](#)
Main subjects: Mathematics ,Physics and Chemistry
- 2012 **12th, Uttar Pradesh Board** [Maharaja Agresen Inter College, Lucknow, Uttar Pradesh, India](#)
Main subjects: Mathematics, Physics, Chemistry.
- 2010 **10th, Uttar Pradesh Board** [S N Public High School, Lucknow, Uttar Pradesh, India](#)

Certifications

- Nov 2017 **Machine Learning** [Global Initiative of Academic Networks\(GIAN\)](#)
At Central University of Rajasthan
Machine learning Techniques with hands-on on Python
With Machine learning Libraries on Sklearn, TensorFlow, Keras, Theano and without Libraries
- Dec 2017 **Machine Learning Application to Internet of Things** [Global Initiative of Academic Networks\(GIAN\)](#)
At IIT (BHU) With Grade-A
Basic Knowledge Of Machine Learning And Theoretical Concepts of IoT
- Jan 2018 **Sketching and Sampling For Big Data Analysis** [Global Initiative of Academic Networks\(GIAN\)](#)
At IIT Kanpur
Basic and Theoretical Knowledge Of Reduce the time complexity of Big Data Algorithms
- Jan 2018 **Big Data Analysis and Mining in Deep Web Repositories** [Global Initiative of Academic Networks\(GIAN\)](#)
At Central University of Rajasthan

Presently Doing

MSc. in Computer Science(Big Data Analytics) [Central University of Rajasthan, Ajmer, Rajasthan, India](#)

Currently I am doing internship in the group of Dr. **Sukant Khurana** at **CSIR-CDRI**.
[at CSIR-CDRI Lucknow](#)

Personal Skills



Places Lived



Languages

Hindi ★★★★★
English ★★★★★★

Other Information

- Hobbies : Exploring new places and things, Travelling, Listening Music, Photography.
- Data of Birth : 07 - March - 1996
- Gender : Male
- Age : 22 years
- Martial Status : Unmarried
- Student member(ID- 30854) of IAI : Institute of Actuaries of India
- Preparing for CT Series of IAI : Institute of Actuaries of India

Other Skills

- Statistics
- Finance and Econometrics
- Data Base Management System (DBMS)
- Value Thinking
- Combinatorial Optimization
- Professional Communication
- Android Studio

Workshop (or internship)

Summer 2018 **Attend The Summer Workshop on Machine Learning Applications to Internet of Things (IoT)** 1st June - 16th July, 2018

Dr. Hari Prabhat Gupta (Workshop Coordinator), Department of Computer Science and Engineering, Indian Institute of Technology (BHU) Varanasi

During this period, I showed extreme dedication and focus on the daily workflow, assignments, and was outstanding in my performance overall. I worked on the following project during my workshop **Apparent motion espial system (AMES) using smart phone**. This 45 days summer workshop is equivalent to the 45 days summer internship or summer training for the same.

Presently

Currently I am doing my internship in the group of Dr. Sukant Khurana at CSIR-CDRI. This is a work- from- home unpaid internship. 14th Nov -

Present

At CSIR-CDRI, Lucknow

Projects

Supply Chain Analytics

In This Project I Work on publicly Available dataset by kaggle For demand Forecasting And by Using SVM(Support Vector Machines) I Predict the Demand Forecast on Python.

Banking Analytics

Presented in VI Rajasthan Science Congress, Oct 13-15,2018

This is case study of Customer Switching Behaviour For Banking Services, In This Project I Create Survey By Using Google Form Which Contain Some 23 Question Related to Customer Satisfaction By The Response of This Question I Generate Some Data And On That Data I use Factor Analysis, Correlation And ANOVA Table for Taking The Conclusion,Done on Python And R.

Health-Care Analytics 1

This is AWS (Amazon web Services) based project, In This Project I Take Publicly Available Dataset Of Heart Disease (Available at UCI). In This Project I use Hadoop Spark To Analysis and Predict. The main Goal of This Project is to Build a model that can Predict The heart Disease Occurrence, Based on a combination of Features (risk factor) Describe the Disease.

Marketing Analytics

In this project I am working on Amazon click stream data and try to get the interest pattern of the user's interest by using an improved Leader clustering Algorithm which is improve by using rough set theory. And also comparing these result with other clustering algorithms like K- medoid and leader clustering algorithm.

Healthcare Analytics 2

In this project I am working on Patient Questionnaire data of the Parkinson's Disease (www.ppmi-info.org/data) and develop prediction models that can classify early PD from healthy normal using machine learning techniques. I get, some result with accuracy score, Logistics regression(Lasso) 95.7919 and SVM(Lasso) 95.5776. So I infer from the study that such diagnostic models might have the potential to be used as an aid in clinical setting by primary physicians when PD expert are not available, for detecting N PD and also for identifying subjects for clinical trials.

References

- Ass. Prof. **Suresh Kumar Choudhary** [+91-9790581558](tel:+91-9790581558)
Department of Data Science and Analytics,Central University of Rajasthan
- Ass. Prof. **Vikas Kumar** [+91-8079044705](tel:+91-8079044705)
Department of Data Science and Analytics,Central University of Rajasthan

November 16th, 2018

Gaurav Dubey