



## RAJ KUMAR, PHD

SEEKING DATA ANALYST OR DATA SCIENTIST POSITION

### PROFILE

As an accomplished academic and materials researcher with over 11 years of experience, I am a proactive and task-oriented professional with a passion for active learning in the fields of Data Science, Machine Learning, and Artificial Intelligence. I specialize in leveraging data insights to enhance customer and business experiences. I am now seeking a transition to a career in Data Science, eager to explore and analyze various data sets, and expand my knowledge of different data domains. I thrive in a collaborative team environment and take pride in effectively communicating data-driven results.

### CONTACTS

	<a href="tel:+919467485015">+91-9467485015</a>
	<a href="mailto:rajkumar.pec@outlook.com">rajkumar.pec@outlook.com</a>
	<a href="https://www.linkedin.com/in/raj-kumar-phd-">https://www.linkedin.com/in/raj-kumar-phd-</a>
	<a href="https://github.com/rajkumarpec">https://github.com/rajkumarpec</a>
	<a href="https://www.researchgate.net/profile/Rajkumar-Kaushik">https://www.researchgate.net/profile/Rajkumar-Kaushik</a>
	<a href="https://scholar.google.com/citations?user=smv1DREAAAJ&amp;hl=en">https://scholar.google.com/citations?user=smv1DREAAAJ&amp;hl=en</a>

## EXPERIENCE

### DATA SCIENCE TRAINEE, ALMABETTER, BENGALURU

AUG 2022-PRESENT

- Possess an extensive understanding of technical skills, including Python, SQL, Machine Learning, and Deep Learning, and have gained hands-on experience using tools such as Tableau, and Power BI.
- Successfully executed numerous projects, including Bike Sharing Demand Prediction, Health Insurance Cross Sell Prediction, and Netflix Movie and TV show, showcasing exceptional analytical and problem-solving skills.

#### Netflix Movie and TV Show-Clustering (02/2022 - 03/2022)

- Developed a clustering model using K-means and hierarchical clustering to segment Netflix's movie and TV show data.
- Utilized NLP techniques, including text cleaning, stemming, stop word removal, and lemmatization to pre-process the data.
- Conducted vectorization of the final text column using TF-IDF and applied PCA for dimensionality reduction.
- Evaluated optimal clusters using the Elbow Method and Silhouette Method.

#### Health Insurance Cross Sell Prediction (01/2023 - 02/2023)

- Developed a Classification Model utilizing advanced algorithms including Logistic Regression, K-nearest neighbour, Random Forest, and XgBoost Classifier to forecast the likelihood of customer interest in Vehicle Insurance.
- Implemented feature encoding, applied SMOTE, and executed Hyperparameter tuning via GridSearchCV to refine model performance.
- Experimented with multiple Performance metrics for various models and identified XgBoost Classifier as the top-performing model, achieving an impressive Recall score of 92%.

#### Bike Sharing Demand Prediction (10/2022 - 11/2022)

- Developed a sophisticated Regression Model utilizing advanced algorithms like Linear Regression, Random Forest, and Gradient Boosting to forecast the hourly demand for bikes, ultimately reducing public waiting times.
- Implemented various processing techniques, including Feature Scaling, Outlier Treatment, Missing Value Imputation, and Multicollinearity analysis, to improve model accuracy and efficiency.
- Employed Hyperparameter Tuning techniques, such as RandomSearchCV, to refine the model and achieved an impressive R2 Score of 86.4% using the XG boost algorithm.

#### Exploratory data analysis of Hotel Booking (09/2022-10/2022)

- Conducted exploratory data analysis (EDA) on Hotel Booking dataset, uncovering insights and patterns in customer behaviour, hotel preferences, and booking cancellations.
- Utilized various techniques including univariate analysis, hotel-wise analysis, booking cancellation analysis, and customer-centred analysis to gain insights into the data.
- Presented findings and recommendations to marketing and business teams to inform decision-making and improve hotel booking strategies.

---

## SKILLS

### TECHNICAL SKILLS

**Programming Languages:** Python, SQL

**Data Analysis:** Pandas, Numpy, Scipy, Matplotlib, Seaborn

**Machine Learning:** SciKit-learn, Tensor Flow, Keras

**Analytics Tools:** Power Bi, Tableau, Advanced Excel

**Database Management:** MySQL, PostgreSQL








**Other Tools:** Git and GitHub, Azure, AWS, Watson cloud and google cloud

### SOFT SKILLS

- Strong performer with consistent results
- Enthusiastic and optimistic approach to work
- Excellent communication and presentation skills
- Proficient in project management techniques
- Strong research orientation and ability to apply research findings to real-world situations
- Skilled in team incubation, fostering collaboration and teamwork

---

## CERTIFICATES

- Full Stack Data Science Program 
- IBM Data Science Professional Certificate 
- Preparing for Google Cloud Certification: Machine Learning Engineer Professional Certificate 
- Intermediate SQL 
- Joining Data in SQL 
- Introduction to Power BI 
- Getting Started with Data Analytics on AWS 

## ACHIEVEMENTS

- GATE (Mechanical Engineering) qualified with an 86.86 Percentile (2012)
- Got MHRD scholarship in Ph.D. and ME

## ASSISTANT PROFESSOR/MMDU, MULLANA, AMBALA, HARYANA

(DEC 2020-JUNE 2022)

- Manage student record during admission and coordinate with DRCC.
- Teach, mentor, and conduct hands-on experiments for B.Tech and M.Tech students.
- Prepare academic reports and supervise examinations as a University Examiner.

## RESEARCH ASSOCIATE, TBRL, DRDO (MINISTRY OF DEFENCE, GOVT. OF INDIA), CANDIGARH

(JULY 2018-JULY 2020)

- Responsible for setup, maintenance, and calibration of Split Hopkinson Pressure Bar (SHPB) to determine material properties under extreme loading conditions in the strain rate regime  $10^2$  to  $10^4$  per second, as well as designing and developing samples and wave shapers to attain stress-strain equilibrium.
- Analysed data using MATLAB, Excel, and Origin to identify trends and patterns, and communicated results through graphs, charts, and figures to support research objectives.
- Responsible for project planning, management, and delivery of reports to various central and government bodies.
- Demonstrated strong attention to detail and accuracy in data analysis, and maintained meticulous records of experimental data and analysis for future reference.
- Collaborated with multidisciplinary teams of scientists, engineers, and professors to achieve project goals and objectives.

## EDUCATION

### DOCTORATE OF PHILOSOPHY (PH.D.), METALLURGICAL AND MATERIALS ENGINEERING

PUNJAB ENGINEERING COLLEGE, (DEEMED TO BE UNIVERSITY), CHANDIGARH, INDIA, (JAN 14 - JULY 2018), CGPA 8.0

### MASTER OF ENGINEERING (M.E), INDUSTRIAL MATERIALS AND METALLURGY

PUNJAB ENGINEERING COLLEGE, (DEEMED TO BE UNIVERSITY), CHANDIGARH, INDIA, (AUG 2011- JULY 2013), CGPA 6.25

### B.TECH, MECHANICAL ENGINEERING

HARYANA ENGINEERING COLLEGE,

KURUKSHETRA UNIVERSITY, (AUG 2004- JULY 2008), 62.5%