

<https://github.com/deepsekhar>[linkedin.com/in/deepsekhar-auddya-857a42104](https://www.linkedin.com/in/deepsekhar-auddya-857a42104)

# DEEPEKSHAR AUDDYA

## Data Scientist/ Data Analytics/Business Analysis

### PROFILE SUMMARY

- An achievement driven professional with 5.11 years of managing projects for data solutions and providing comprehensive statistical & analytical models
- Act as a leader with vision for the use of data and analytics to drive critical business decisions, business strategy and product management activities
- Ability to take ownership and initiative, to negotiate, influence and build consensus and successfully navigate within a data driven environment
- Performed exploratory and targeted data analyses through the systematic application of statistics and mathematics
- Experience in identifying data needs, drive methods and standards for data prep, data quality and governance.
- Expertise in demonstrable analytical skills with an emphasis on statistical modeling, decision science and predictive analytics
- Provides influence and technical advice to Jr. level team members when developing solutions
- Understand the purpose of work and create workflows, project phases and draw inference of that
- Design and build modern machine learning and statistical methods or finding structure in disparate datasets
- Persuasive and articulate in presenting complex ideas.

### Core Competencies

- Data Scientist
- Consulting
- Business Analysis
- Machine Learning
- Visualization
- Automation

### SKILLS

#### Data Science and Data Analysis

Correlation, multicollinearity, EDA, Anova, Hypothesis Testing, Probability Distribution, MS-Excel, MSBI

#### Machine Learning

Supervised and Unsupervised Learning

#### Deep Learning

ANN, DCGAN, CNN, RNN, LSTM, Tensorflow, Keras

#### Natural Language Processing

Bag of words model, seq-seq model, Text classification, Word Embedding, Text Summarization, Topic Modeling, LDA, LSA, NER, Sentimental Analysis

#### Python

Python programming, Numpy, Pandas, Scikit-Learn, NLTK, Spacy

#### R

R Programming, data. table, data frame, data matrix, Dplyr, lapply, udpipe, glm, tm

#### Alteryx

Crew Macro, Macros, Multi row, multi field, multi-filed binning, append, Union, Fuzzy match, regex, date-time, In-DB tools

#### SQL

Multiple Joins, Union, Store Procedures, Functions, Triggers, Views, Cursors

#### SAS

PROC UNIVARIATE, PROC SQL, PROC FREQ, PROC MEANS, PROC ANOVA, GLM and mixed models

#### Tableau

Conditional Formatting, Sets, Parameters, REGEXP, Analysis reference, Dashboard Actions, LOD

## EMPLOYMENT DETAILS

**Senior Data Scientist**  
**Sight Spectrum - Bangalore, Karnataka**

**Feb'20 to Present**

**Job Responsibilities:**

- Develop algorithms and build data models for use in applications and process improvements
- Building and productionizing predictive models on large datasets by utilizing advanced statistical modeling, machine learning, or other data mining techniques
- Building models on text data using advanced word embedding and attention techniques
- Creating text classification model

**Senior Data Scientist**  
**Accenture – Bangalore, Karnataka**

**Job Responsibilities:**

**Mar'19 to Dec'19**

- Selecting features, building and optimizing classifiers using machine learning techniques
- Enhancing data collection procedures to include information that is relevant for building analytic systems
- Provided training on ML and Deep Learning among the team members
- Doing ad-hoc analysis and presenting results in a clear manner
- Automating models and metrics that track product performance, ensure data quality and integrity
- Utilizing predictive models to build strategies that can address key issues in growth and operations
- Bring in quality/process improvement expertise in the realm of data science
- Conduct code reviews on components and applications to ensure adherence to the development standards and best practices in Python
- Contributes to design and prototyping of medium to high complexity machine learning systems
- Providing thought leadership in statistics and data analysis to the entire analytics organization

**PREVIOUS EXPERIENCE:**

**Data Analytics/Data Scientist**  
**PricewaterhouseCoopers Service Delivery Centre (Kolkata) Private Limited – Kolkata, West Bengal**

**Sep'16 – Feb'19**

**Job Responsibility:**

- Design, conduct, and report results from prototype or proof-of-concept research projects that focus on 1) new tools, methods, or algorithms, 2) new scientific domains or application areas, or 3) new data sets or sources
- Analyzing large data sets, including synthesizing quantitative results, deriving business implications and making clear recommendations
- Work independently, develop prototype concepts and manage experiments to create data-driven solutions to business problems
- Client coordination single handedly and Prepared power point presentation and explain it to the client to get work from them
- Independently reason and understand models in general from and probabilistic perspective
- Devise and utilize algorithms and models to mine big data stores, perform data and error analysis to improve models, and clean and validate data for uniformity and accuracy
- To help develop and support the use of the Analytics Centre of Excellence policies, tools, and best practices across the enterprise
- Efficiently manage time for timely project deadlines as well as ad hoc requests
- Helping the team by sharing experiences and knowledge Liaising with team to get the best delivered product to the client and creating dynamic solution of a particular problem and largely engaged with the client and consultants
- Managing data mining and data modelling in coordination with the finance department

**Roles & Responsibility:**

- Understand and analyses existing process to determine futuristic solutions. Assist clients in identifying the right sourcing strategies. Testing and developing the tools for the Automation and Analytics.
- Pending analysis to reduce the pend customers solution towards monitory benefits.
- People management skill building effective teams to complete tasks.
- Root cause analysis of different exceptions to reduce the exception generation volume.
- Preparing KPI parameters and Dashboard for the client. Perform analysis, draw conclusions and formulate strategic recommendations. Customer Interaction Analysis: Customer sentiment analysis to reduce customer complaints
- Analyzing query of the existing system and come up with solutions.
- Create process re-engineering approach/roadmap journey for the process.
- Lead client workshops/ discussions for the Business Values.

**NOTABLE ACHIEVEMENTS ACROSS CAREER**

- Successfully organized Basic and Advance Excel Training
- Designed and executed end to end workflow by auto; praised by client for efforts
- Acknowledged by management for attention to detail and high potential for promotion
- Redesigned project plan and strategy after encountered problems; recognized for ability to turn it around and finalize it on time
- Provided trainings in ML, Python, R, Alteryx
- Bagged:
  - Spot Award in 2017 year
  - Star Award in organization level in 2017 year
  - Innovation Award in 2018 year
- Organized VBA and Data Model training and KT Sessions

**CERTIFICATIONS**

- Business Analytics and Business intelligence from Orange tree Global (SQL, SAS, R, Tableau, VBA)
- Alteryx Certification (Core & Advanced) from Alteryx Community
- Business Analysis and Data Science from IIT KGP (Python, R)
- Fundamentals of Visualization with Tableau by University of California, Davis (Coursera)
- Introduction to Data Analytics for Business by University of Colorado Boulder (Coursera)
- Predictive Modelling and Analytics by University of Colorado Boulder (Coursera)
- Machine Learning by Stanford University (Coursera)
- SQL for Data Science by University of California (Coursera)
- Tensorflow for Deep Learning with Python (Udemy)
- Deep Learning Artificial Neural Network from (Udemy)
- Design Thinking (Jobs to be done) by PwC L&D team
- Design Thinking approach and implementation by Accenture LKM team

**Professional Trainings**

- SAS BASE
- SAS ADVANCED
- DATA ANALYTICS and SAS 9

**BUSINESS SKILLS**

- TTT Certified: Certified as a trainer by the PwC L&D initiative
- Business Communication Skills and Leading Teams: As part of the Learning & Development initiative
- Certified in business communications, business emails, critical thinking and team leading capabilities.

## EXTRACURRICULAR ACTIVITIES

- Actively involved in organizing blood donation camps, coordinating for blood donation during emergencies and organized cultural programs
- Attained Entrepreneurship Success award as a part of Institutional Development
- Participated & won in many essay and elocution competition
- Organized various fest & events in college & clubs

## EDUCATION

Bachelor in Commerce, Related subjects Economics and Statistics

## Projects and Case Studies

### **Title: Project on Credit Risk Modelling**

**Description:** Managed project related to financial markets wherein examined the trust worthiness of a prospective customer and his/her possibility of defaulting on loan

### **Title: Project on Churn Analysis**

**Description:** Predict whether the customer will leave the bank or not using Artificial Neural Network by creating a classification model

### **Title: Project on Credit Risk Modelling**

**Description:** The task was to build a Behavioral Credit Risk Model based on a large sample of data. The data had various information related to behavior transactional details of the customers and their performance in repayment.

### **Title: Project to predict a class of a new feature**

**Description:** Use KNN to create a model that directly predicts a class for a new data point based on the features

### **Title: Project to predict home sales price**

**Description:** Predict home sales price with the xgboost algorithm

### **Title: Project on Medical Science**

**Description:** Identify the person has disease of pneumonia or not by watching an x-ray Using VGG16 transfer learning from Keras and in output layer used sigmoid activation function. Model compile done by loss categorical cross entropy with Adam optimizer

### **Title: Project on Stock-Price-Prediction**

**Description:** Project related to predict the future month stock on the basis of past two years data using-Keras-and-Recurrent-Neural-Network. Took 60-time stamps to compute the output

### **Title: Project on Customer Satisfaction Survey**

**Description:** Project related to retail domain based on Customer Satisfaction Survey where the objective was to find out the key factors influencing the overall customer satisfaction using OLS Regression.

### **Title: Project on Time series Forecasting base on Sales data**

**Description:** Project related to sales of an organization where future sales could be forecasted using the various techniques of smoothing out the random fluctuations.

### **Title: Project on Testing of Hypothesis on Clinical data**

**Description:** Steered case study about Blood Pressure of different respondent pertaining to Clinical Research to check whether the average blood pressure in the population was equal to hypothesized value or not.

**Title: Project on Sentimental Analysis**

**Description:** Project related to analyse the social sentiment of the client. Had to create API to extract all the comments and replies from Client's Facebook, Twitter page and website. Build a word cloud to know which word is emphasizing more in comments. Shared work effectiveness and effort analysis. Showed them in which aspect they need to focus

**Title: Project on Due diligence of different clients in various sector like retail, healthcare, banking etc.**

**Description:** Build data model according to the data set in Alteryx. Use SSAS and build cube deployment and SSMS to store and reconciliation. Build dashboard from every aspect in Tableau to show portfolio matrix, monthly charts, scatter charts and draw infer of that. The purpose was to show with different parameter and metrics of business how they should focus and which area to focus before due diligence.

**Title: Project on Automatic Text Summarization from a large amount of Text and 10- K reports to identify the risk factors**

**Description:** Within R Shiny build dashboard and automated the particular link and href with the help of ticker and **Edgar R**. Use lexicon rank algorithm to summarize the text as an adaptive learning. Prepared N-Gram and word network graph to identify the co- occurrence and relation of the words

**Title: Project on automation of data population from a pdf with their different state wise tax return.**

**Description:** The data structure was numerous unstructured for each state. We have used multiple for loops and perform various data cleaning and manipulation steps to identify the range of page of a particular state. Then to identify the exact value of the columns we used general regular expression and other functions of R.

**Title: Project on analyze the company response on the basis of customer complaint for a bank.**

**Description:** There were multiple complaints as text document. We create a doc as corpus use cleaning and preprocessing to remove unnecessary words. Then we use topic modelling to cluster words on their similarity and presence. We have used LDA and LSA model to fit in the data.

**Title: Project on analyze patient exposure**

**Description:** There were multiple patient's records with the use of drugs. We created dashboard in R Shiny usage of drugs along with their age, use rate, region cumulatively. It was an automated dashboard with the data source inside. We use linear model algorithm for all region if the new data comes in the use of drugs has to be liner or upward trend

**Title: Confidential\_Nexus\_R integration**

**Description:** Various calculations performed with creation of advance functions and iterating loops with the string from the file name. Prepared an MDM sheet from multiple excel files with various calculation technique which will allow to integrate to the existing server