KOUSHIK KHAN

IBM India Private Limited, Embassy Golf Links, Bangalore - 560071, India

Phone: (+91) 7501-923-251 • **Email:** koushikkhan38@gmail.com • **LinkedIn:** https://www.linkedin.com/in/koushikkhan/

ABOUT

I am currently working as a Data Scientist at IBM, having five years of industry experience in the field of advanced analytics and machine learning. I completed my graduation and post-graduation in Mathematical and Computational Statistics from Visva-Bharati University, Santiniketan, West Bengal, India.

I take a special interest in *Linear Algebra*, *General and Statistical Programming* and *Technical Writing*. My Personal web page is maintained over here.

EDUCATION

Visva Bharati University, Santiniketan, India *June, 2013 - June, 2015*
Master in Statistics
Department of Statistics
Thesis Title: Joint Modeling of Longitudinal and Time to Event Data

Visva Bharati University, Santiniketan, India *July, 2010 - May, 2013*
Bachelor in Statistics
Department of Statistics
Thesis Title: Method of Estimation, A Generalized Approach

KEY SKILLS

- **Programming**
  - Object Oriented Programming, R, Python, SQL
- **Predictive Modeling and Machine Learning with structured and unstructured data**
  - Regression and Classification (Linear and multiple linear and non-linear regression, k-Nearest Neighbor, Naive Bayes, SVM, Tree based modeling, Perceptron Models), Cluster Analysis & Topic Modeling, Marketing Mix Model, Forecasting using Recurrent Neural Network (RNN, LSTM, GRU), Natural Language Processing
- **Software and APIs**
  - Database applications (SQL and No-SQL), Deep Learning APIs (Keras), Dashboard (R Shiny), Deployment tools (Python REST APIs and Docker)
- **Cloud Platforms**
  - IBM Watson Studio for ML model development, Microsoft Azure

PROFESSIONAL EXPERIENCES

2016, Oct. - Present	IBM India Pvt. Ltd.	Data Scientist
2015, Oct. - 2016, Aug.	BCausE Enterprise Pvt. Ltd.	Analyst

PROJECT EXPERIENCES

- **Optimization of ROI on Advertisement Spend Investment - External Client for IBM**

  - Media Industry deals with several key business metrics like *sales volume* of a certain commodity because of an advertisement, *likely to buy* score for different commodities etc. These KPIs are often required to be forecasted for efficient broadcast of advertisements into multiple advertising slots on a specific day. Efficient broadcast helps to promote products or commodities in a better way
  - The core business problem was to build forecasting models for the aforesaid KPIs using deep learning
  - Developed data transformation pipeline and forecasting models using *LSTMs* and *GRUs*
  - Based on forecasted values ROI metric has been computed along with the importance scores of various features available
  - Worked on the integration of the models with the backend UX platform

- **Marketing Mix Modeling - External Client for IBM**
  - Developed Marketing Mix Models and Optimizers for different brands of Nestle in various markets across the globe
  - The models were developed to understand the sales behavior, analyse base-incremental sales volumes and calculate the return on investments of key marketing and promotional factors like TV, Print Media and Online Awareness Programs on FaceBook / YouTube on Nestle products
  - The R language has been used extensively for data wrangling, modeling and later on models have been deployed in Microsoft Azure Platform using REST APIs
  - Developed data visualization and simulation (explaining what-if scenarios) tools in R Shiny web framework
- **Commodity Hedging Price Forecasting - External Client for IBM**
  - Developed Recurrent Neural Network model for time series data for daily hedging prices for several crops like corn, wheat etc.
  - Worked on improvement of the existing models by analyzing new data collected from the agencies like Chicago Board of Trade (CBOT)
  - Worked on deployment of the models in Microsoft Azure Cloud Platform using Azure Databricks services
- **Small Scale Product Recommendation Engine - External Client for IBM**
  - Developed a small scale product recommendation engine for the Vietnam market of the client
  - User (retailer) specific sales information as well as product specific attributes were provided by the Vietnam marketing team
  - The idea was to build a system to identify most likely products to be recommended based on the purchase histories of the retailers and product-to-product similarities
  - Contributed in model deployment by leading the off-shore team
- **Contextual Information Retrieval System - External Client for IBM**
  - Developed a small scale IR system based on the concept of latent semantic indexing to cater relevant answers to queries for FAQ
- **Cognitive Test Automation - IBM GBS Productline, IBM**
  - Developed an automated ticket (text) classification system as part of the core cognitive engine
  - Developed automation tools, wrappers and REST services for NoSQL database integration, data validation, data processing, model maintenance and retraining etc.
- **Dataflow simulator design and Analysis, BCausE**

– Contributed in developing a simulator to compute the likelihood of data flow among several wireless network devices like multiple GSM SIM slots and wi-fi module for an electronic device

AWARDS & RECOGNITIONS

2016 Granted DST Fellowship for excellent performance in Masters Programme by DST, Govt. of India
2015 Achieved Second position in M.Sc. Examination
2013 Achieved First position in B.Sc. Examination
2011 Achieved Scholarship for Higher Education by Govt. of India

OTHER JOB RELATED ACTIVITIES

Technical Articles:

- A Not So Short Introduction to Object Oriented Programming using R
- Demonstration of Central Limit Theorem with Simulated Binomial Variates
- Effect of BoxCox Transformation on Non-Normal Data Sets
- Principal Component Analysis – An Introduction with R Implementation
- Basic Data Manipulation with dplyr
- Testing of Hypotheses: A Quick Refresher

Personal Project Notebook:

- Colab - Stock Price Forecasting using Recurrent Neural Network