

Contact

91-9751184410 (Mobile)
ashishbt08b004@gmail.com

www.linkedin.com/in/ashishk64
(LinkedIn)
www.facebook.com/profile.php
(Other)

Top Skills

Matlab
Python
Microsoft Excel

Languages

Hindi (Native or Bilingual)
Bhojpuri (Native or Bilingual)
German (Limited Working)

Certifications

Neural Networks and Deep Learning
Improving Deep Neural Networks
An entire MBA in 1 course
DevOps for Data Scientists
Azure Machine Learning
Development: 2 Learning ML Studio

Honors-Awards

1st Prize, Best Paper and
Presentation
2nd Prize, Best Essay
Full tuition waiver
Merit Recognition Award

Publications

Data Science Page
Making India the great nation that it
is
Everything Analytical
Learning Predictive Analytics with
Python
Gorkhaland: Salvation of Identity
First, Development Second

Ashish Kumar

Principal Data Scientist | IITM | YIF | PACKT
Chennai

Summary

~Data Science & Machine Learning professional with 7+ years of work-ex (Valid US B1 Visa)
~ Github - <https://github.com/ashishbt08b004/Experiments>
~Author, Learning Predictive Analytics with Python. Author, Mastering Pandas. Both published by PACKT.
~ Rich hands-on and project management experience (25+ projects) in creating Machine Learning solutions.
~ Versatile (75+ leads) pre-sales, solutions architecture, PoC development and data science consulting experience.
~ Lucid Communication of complex data science concepts using both verbal and written medium.
~ Experienced in conducting webinars and training sessions.
~BTech, IIT Madras |Young India Fellow 2012-13, Ashoka University
~ Financial Modeling| Business pitch expert

Programming Language/Tools : [Python, R, SQL, R Shiny, SAS, Julia, Go], [Tableau, Power BI, Excel, Google Analytics], [Hive, PostgreSQL, Redshift, Elastic Search, Spark], [Azure, AWS, Docker], [JIRA, Confluence]

Algorithms : Linear & Logistic Regression; Survival Analysis, DBSCAN, k-Means; kNN; Naïve Bayes; SVM; Shapelets; Latent Dirichlet Allocation, Linear Discriminant Analysis, Latent Semantic Analysis, Hypothesis Testing, ANOVA, RFM, BTYD, Apriori, Voronoi, Geospatial Analytics, Tf-IDF/Word2Vec, CNN, RNN

Business Domains: Manufacturing, FinTech, EdTech, Transportation & Logistics, Healthcare, Pharma, Retail, E-Commerce, Big Data Security, Urban Mobility.

Presentation: Communicating ML results & business insights by creating dashboards/analytics products/report in R Shiny, Tableau, Power BI, Jupyter Notebook, MS-Excel, PowerPoint etc.

What excites me:

- ~ Working on challenging and intellectually stimulating business problems to solve through data science
- ~ Working with emotionally intelligent people who can think beyond money & themselves
- ~ Perpetual learning. Student for life.
- ~ Debate over data, public policies, philosophy, pop culture

Experience

teX.ai

Principal Data Scientist

2019 - Present (1 year)

Leading R & D and Product Development of teX.ai.

teX.ai has been recognised as 'Top 25 ML startups to watch out for in 2020 by Forbes.'

Indium Software

4 years

Principal Data Scientist

2019 - Present (1 year)

- Spearheading the R&D and Product Management of the cutting-edge Text Analytics product offering Text Extraction, Summarization and Classification for enterprises. Get in touch to know more.

Text Extraction - Tabular and Peripheral data from PDFs, Images, Word Docs and Websites

Text Summarization - Latent Topic Modeling, Keyphrase Detection, Custom Named Entity Recognition, Text Matching and Clustering

Text Classification - Classify documents in categories for better organisation.

Senior Manager, Analytics

2018 - 2019 (1 year)

- Project Management for 5 running projects and pre-sales for qualified leads.
- Grew a large manufacturing account 4X in terms of team and ticket size.
- ~100% project extension success rate.
- Pre-sales; Requirement-Capability matching by talking to US/Europe clients; Contributing to Scope/SoW of projects; Deciding the team structure.

- Architecting the analytics solutions and then implementing them with the teams.

Abnormal Chamber Detection & Predictive Maintenance - Led development of IoT based analytics product to detect abnormal chambers for a semiconductor wafer manufacturer in CA, USA. Algorithm & tools - Mahalanobis distance, Hotelling T2, R-Shiny

Product Categorisation of 50mn+ products - Led categorization of 45mn+ products using text classification ML algorithms improving search results for an e-commerce firm in SF,US. Deployed the solution as an API. Algorithms & tools - Naive Bayes, SVM, Hierarchical Classification, Python, AWS.

Dropout/Failure/Withdrawl prediction - Implemented Dropout/Failure prediction for students from a classroom course for a learning management company in US. Algorithm & tools - Xgboost, R.

Intelligence Layer to knowledge on web - Led the development of NLP and text analytics based solutions – topic map for documents, document/publisher similarity, topic classification, entity recognition etc for a knowledge solutions company in Massachussets, US. Algorithms & tools - Latent Dirichlet Allocation, Nmaed Entity Recognition, Neural Networks, Python, D3.js.

Market & Product Analytics for a Fintech company - Conducted product/market analytics for a Fintech firm in SF,USA and devised growth strategies for the firm. Algorithms & tools - RFM Analysis, k-means/DBSCAN clustering, A/B Testing, R, Python, Periscope

Customer personas for an Event Management Company - Clustering customers into groups with a marketable personas for better targeted marketing. ALgorithms & tools - k-means, DBSCAN, tSNE, Expectation Maximization, R, R Shiny.

Program Manager, Analytics
2016 - 2018 (2 years)

Overview: Part of the senior management team driving sales, digital marketing, strategy and delivery of the firm. Looked after 10+ analytics engagements, devised and implemented various algorithms to help clients. Led research & PoC for 10+ prospective projects. Played active role in 25+ pre-sales

conversations. Groomed and managed the 30+ membered Data Science team. Wrote blogs.

- Geospatial analytics for the largest taxi-ride provider company in South East Asia| Tools & Techniques: Hive, R Shiny, Leaflets, Haversine formula.

- Automated Hive Query Generation for a Big Data security product company based in San Francisco Bay area| Tools & Techniques: Hive, Impala, SparkQL, R Studio, Apriori

- Text based entity mapping and search for Manhattan-based BI provider| Tools & Techniques: Python NLTK, Levenshtein distance, Elastic Search

- Historical Price Analytics for products for a price-comparison e-commerce website based in San Francisco|| Tools & Techniques: Elastic Search, Kibana, Timelion

- Growth Analytics for a price-comparison e-commerce website based in San Francisco|| Tools & Techniques: Google Analytics, Google Tag Manager, Mixpanel

- Electricity demand forecasting based on weather data using Generalised Additive Models for a clean energy client based in NY| Tools & Techniques: R Studio, Time Series Analysis

- Detection of defective pieces using power consumption time-series data| Tools & Techniques: R Studio, kNN, Azure ML Studio, Python

- Data Science Training for employees of national statistical agency in UK| Tools & Techniques: Python, NLTK

Blogs:

- 1) <http://www.noahdatatech.com/leveraging-your-gps-data-using-geospatial-analytics/>

- 2) <http://www.noahdatatech.com/iot-analytics>

Great Learning
Course Mentor
2020 - Present (less than a year)

Mentoring students on their Data Science course.

Packt

Author, Mastering Pandas

2019 - 2019 (less than a year)

Simplilearn

Trainer & SME

2019 - 2019 (less than a year)

Trained a batch of 40 students in a virtual classroom on Advanced Machine Learning concepts using Python as the implementation tool.

Trainer NPS in top 5% percentile. Got offered to join the premium pool of trainers at Simplilearn.

Packt

Author, Learning Predictive Analytics with Python

June 2015 - March 2016 (10 months)

Chennai Area, India

- Wrote a 350-pager/9 chapter book demonstrating the concepts of predictive analytics with Python emphasizing on Data Cleaning, Wrangling, Modeling, Validation, and Visualization. Used publicly available datasets to develop original content
- Liaised with a team of editors, reviewers, experts, designers etc. Got an offer to write the next title from the publisher.
- Rated must-read for data science enthusiasts by Analytics India Magazine . Garnered positive reviews on Amazon 50K USD in revenue. 300+ copies sold in the first 3 months. Selected for display and early sale at the PACKT publication website during the Python week.
- Got an offer to write the next title from the publisher. Selected to become part of a PACKT video course.

Tiger Analytics

Senior Analyst

May 2014 - 2016 (2 years)

Chennai Area, India

Project 1: Wheel failure forecasting for a leading railroad car pooling client

Developed a SAS model to forecast wheel failures using Survival Analysis on historical failure data of 30 years from a leading railroad car pooling client. The model could predict the number of failures with an accuracy of around 1.5-2%.

Developed a tool with VBA-based MS Excel interface to visualise and summarize the results of the aforementioned model. This tool is used by the client to plan quarterly spend on maintenance & repair.

Tools/Techniques used: SAS, VBA, R; Survival Analysis, Monte Carlo Simulations, Seasonality

Project 2: Merchant, Customer and Sales analytics for an Online Payments client

Integrated telesales data of a newly acquired firm with the Salesforce data of the client in Teradata using Python

Analysed transactions data from the client's customers, built Tableau dashboards and a Python simulation to notify the customers of their operational and financial status

Tools/ Techniques used: Teradata/SQL, Python, Tableau

Project 3: Impact assessment of a packaging change on sales for a Pharmaceutical manufacturer

Assessed the impact of a new packaging method on drug sales for 100+ SKUs
Calculated the lift in drug sales and contribution of drug sales to the total sales attributed to the new packaging

Project 4: Marketing-mix modeling and RoI calculations for a retail Medicare product manufacturer

Assessed the contribution of 30+ marketing stimuli on sales. Calculated the Rols of the various marketing stimuli

Project 5: Content development for training module on Data Science with R

Created and validated the assignments by solving them for a couple of chapters in the training module

Project 6: Time Series model automation for for a leading US based railroad car pooling client

Implemented ARIMA model for cost forecasting for a railroad car company
Wrote an algorithm to automate the ARIMA implementation to find out the optimum ARIMA parameters

ZEEF

Curator, Data Science Page

October 2015 - December 2015 (3 months)

Chennai Area, India

Curated a peer-reviewed page for data science containing list of online material needed to self-train oneself as a Data Scientist.

The page has 100+ links, 2.5K+ views, 600+ clicks and has gained significant traction in the data science community.

SughaVazhvu Healthcare

Business Analyst

June 2013 - May 2014 (1 year)

Thanjavur Area, India

- Designed the Excel-based interactive MIS for SV. This MIS became indispensable tool for performance review and management across the network. Analysed the MIS data on Excel on a monthly basis for performance management and business insights.
- Created detailed financial model and funding proposal for a Mobile clinic and semi-urban clinic. The model garnered a substantial funding for SV.
- Led the efforts to design (interior and exterior), drafting the operational plan, negotiating with vendors and on-the ground execution of the Mobile clinic on the ground.
- Designed interactive Open Data Kit XML-based survey forms for data collection of Cardio Vascular Diseases Risk factors. Led the implementation of the activity. Analysed monthly data to target at-risk patients and manage agent performances.
- Designed the Excel-based interactive MIS for SV. Standardised the SQL queries and data-parsing algorithms needed to update the MIS. Analysed the MIS data on Excel on a monthly basis for performance management and business insights.
- Got selected to talk about SV's disruptive healthcare model at a social entrepreneur summit in Nairobi, Kenya organized by Ashoka foundation

Ashoka University

Young India Fellow

May 2012 - May 2013 (1 year 1 month)

Young India Fellowship is a one year, multi-disciplinary, (Liberal Arts +Leadership) postgraduate flagship programme, run by Ashoka University. Got a full scholarship of Rs.5.5 lakhs to pursue the program

Was one among the 97 Fellows selected out of 3000 applicants in this class, on full scholarship. I studied Group Dynamics, Leadership, Business, Statistics, Arts Appreciation, Sociology, Philosophy etc.

Worked with British High Commission for mapping India's FDI outflow footprint.

Wrote papers on similarities between ideals of Gandhi and Islam, weddings in Bihar, reimagination of Shakespeare's Tempest in the context of Naxalism in India

Interviewed Madhubani painter Bharti Dayal and presented a report on Madhubani painting

Made a movie on an NGO (Chintan) working with ragpickers in Delhi.

Indian Institute of Technology of Madras

Project Associate

July 2008 - May 2012 (3 years 11 months)

Chennai

Worked on a industry sponsored project by Caterpillar with Mathematics department of IITM.

Worked on a Game Theory project with a Department of Management Studies professor.

Worked on a project on Mutual Fund performance analysis in Indian markets with a Department of Management Studies professor.

Energy Alternatives India/Oilgae

Research Analyst

December 2010 - January 2011 (2 months)

Chennai

Made a report on feasibility of a certain strain of algae to produce biodiesel.

The report became a part of the final submission made to NTPC for alternative energy sources.

National Centre for Biological Sciences

Research Associate

May 2010 - July 2010 (3 months)

Bangalore

A project in developmental genetics.

Created a genetic screen to identify responsible Transcription Factors (TF) for sensillae development. Identified 4 such TFs out of 20.

Designed the project flow of the experiment, scheduled the crosses of the insects and prepared the antennae glass-slides.

Found 2 TFs playing important roles in antennae development.

Worked under the guidance of famous scientist Dr. Veronica Rodrigues.

Education

Indian Institute of Technology, Madras

B.Tech, Operations Research/Biotechnology · (2008 - 2012)

Ashoka University

Young India Fellowship, Post Graduate Certificate in Liberal Arts and
Studies · (2012 - 2013)

Edvancer Eduventures Pvt. Ltd.

Certified Business Analytics Professional, Business Statistics · (2013 - 2014)

Kendriya Vidyalaya

High School, Science