ABHINAV KUMAR

Phone: +91 9611050464 | Email: a.kumar2102@gmail.com

Experienced Data Scientist with over 8 years of industry experience and a demonstrated history of working in the Product and services industry. Skilled in Python, Web Scraping, Machine Learning, Deep learning, Statistical Data Analysis, Data Visualization and Predictive Modelling. Strong Data Professional with Expertise in leading & delivery of several Analytics, data science and deep learning projects.

CORE COMPETENCIES

Strategic Thinking: Able to influence the strategic direction of the company by identifying opportunities in large, rich data sets and creating and implementing data driven strategies that fuel growth including revenue and profits.

Data Collection: Understanding the data requirement, identifying the data variables affecting the outcome, identifying the data sources and collecting relevant data from the client and from the web through using web scraping tools.

Modeling: Design and implement statistical / predictive models and cutting-edge algorithms utilizing diverse sources of data to predict Sales Pattern and do cost and profit analysis for marketing campaigns

Analytics: Utilize analytical applications like R to identify trends and relationships between different pieces of data, draw appropriate conclusions and translate analytical findings into risk management and marketing strategies that drive value.

Data Visualization: Develop tools and reports that help users' access and analyze data resulting in higher revenues and margins and a better customer experience.

Communications and Project Management: Capable of turning dry analysis into an exciting story that influences the direction of the business and communicating with diverse teams to take a project from start to finish. Collaborate with product teams to develop and support our internal data platform and to support ongoing analysis.

SKILLS and TOOLS

≈ **Languages**
   Python, R, SQL
≈ **Statistical Packages**
   SciPy, NumPy, Pandas, matplotlib, Scikit, NLTK, Scrapy, TDL, VBA, Selenium
≈ **Software**
   MS Excel, Tableau, GIT (Version control)
≈ **Statistical Model**
   NLP, Word2Vec, Understanding of Supervised and Un-supervised learning, Linear and Logistic Regression, Bayesian Classifiers, Clustering, Support vector Machines, Neural Networks, Decision Trees, Ensemble Methods, (Random Forest, Boosting Trees), ANOVA and Basic Statistics.

EXPERIENCE

Velkon Analytics Pvt. Ltd.   **Bangalore, India**
Software Engineer   **Mar2018-Till Date**

- ≈ Handled several web scrapings, Data analytics and Data Visualization projects from end to end.
- ≈ Built scrapers with various tools as per the requirement. Worked with Request, Beautiful soup, Scrapy, Splash, selenium Packages.
- ≈ Worked on the automation of the scraping scripts.
- ≈ Built scraper to scrape Job profiles from Job search sites like LinkedIn, Naukri based on Geo location and Skill, scraper to collect 10 million reviews of products from ecommerce websites and scraper to get real time flight prices across various websites, with IP and header rotation to avoid getting the IP blocked.
- ≈ Developed Scrapers with IP and header rotation, to avoid getting blocked.
- ≈ Developed Tableau Dashboard for a major hotel aggregator.
- ≈ Developed NLP algorithm to map Job requirements posted on the website to O*net Code for a Major Job portal
- ≈ Developed Machine learning algorithm to identify "Intent to switch" for an International Job Portal.
- ≈ Developed Machine learning algorithm to predict the time when there is risk of driver stock out for a food delivery platform.

Tally Solutions India Pvt. Ltd.   **Bangalore, India**
Software Engineer   **June 2011-Feb 2018**

- ≈ Worked at Tally Solutions Private Ltd., R&D from June, 2011 till Feb, 2018.
- ≈ Assist in developing in-built analytic capabilities in the upcoming Product.
- ≈ Assist in developing internal tools for data analysis
- ≈ Perform quantitative analysis of campaigning and sales trends to recommend Marketing Strategies.
- ≈ Involved in business and functional requirement gathering, creating design documents, development and Bug fixes as per SDLC.
- ≈ Involved in enhancing the business engine for Tally ERP Product using C++.
- ≈ Played a key role in developing various modules of Tally.ERP9 like Retail, Banking and Taxation.
- ≈ Participating in discussions with peers to analyze the business requirements that need to be integrated in the product and working for the same.

PROJECT DETAILS

Velkon Analytics Pvt. Ltd. (March 2018-Till Date)

1. Created machine learning algorithm to classify the jobs posted on the website into one O*net Job Description for a major job search portal.

   Overview: O*net (occupational information network) is a list of hundreds of occupational definitions maintained by US Department of labor, to help the job seekers understand today's world of work in United States. The job portal needed to classify every job posted on its website into most suitable O*net description based on the job requirement, in order to utilize this efficiently for connecting right candidates with right job requirement.

Technology used: Python, NLTK, Word2Vec, TFIDF

Responsibility: Communicating with client to understand the requirement and building the algorithm to classify the job requirements. Involved cleaning up the data, Noise removal, Normalization, stemming to get the roots of every word, feature extraction, Topic modelling, creating word2vec model with TFIDF weighing and then using measuring spatial distance if every job with each O*net classification to find the list distant O*net title.

2. Developed Scrapy web scrapper to collect 10 Million reviews of different items from across ecommerce platforms.

   Overview: The client wanted reviews of a list of electronic items across various ecommerce websites.

   Technology used: Python, Scrapy, Splash, Request

   Responsibility: Involved in client requirement understanding and delivery of the project end to end.

3. Created Tableau Dashboards for a major Hotel aggregator.

   Overview: The client needed a dashboard where he can clearly visualize the price variation per hotel for a particular check-in checkout data, on a any given day and time.

   Technology used: Tableau

   Responsibility: Was responsible for the development and delivery of the dashboard as per the client requirement and doing changes in it as when is suggested by the client.

Tally Solutions India Pvt. Ltd. (June 2011-Feb 2018)

4. Created a Knn classifier model to group the channel partners in different categories based on their engagement behavior which led to formulation of a new Marketing/support Campaign Strategy and resulted in a 13% increment in the quarterly sales of the company.

   Overview: Tally handles its sales and service across the country with a huge partner network of 20000+. And to enable them for service and help them with their sales and marketing, there is a huge cost and effort that goes into it. We ran various level of analysis to cluster these partners into various categories on various levels (e.g. sales partner/service partner, Based on Total Sales and level of Activeness and engagement with Tally, and others), which helped in directing the right kind and level of support to different group of partners.

   Technology used: Python, K-means, Knn, Naive Bayes and Logistic Classifier

   Responsibility: Was part of the team responsible for seeking the right Data and running the analysis end to end.

5. Developed a web scrapper using Scrapy, Beautifulsoup and Request Packages of python.

   Overview: Tally ran an internal project to develop an analytics tool which will help the businesses using Tally understand how they are performing compared to their competitors in the market who are also running Tally.

   Technology used: Python, Scrapy

   Responsibility: Was responsible for understanding the requirement and developing the scrapper and doing Adhoc changes in the scrapper from time to time for different websites as per the project need.

6. Created Adhoc Sales reports and different analysis reports in Excel while working on marketing campaigns to give a quick insight into daily performance of the Campaign.

   Overview: Different Marketing and Sales Campaign are run from time to time for a duration varying from a month to a Year. During these campaigns several sessions are held across different parts of the country and quick overviews are required if the campaigns and sessions were helpful and successful. This needed quick generation of a reporting dashboards which will give a proper insight to the sales and marketing team.

   Technology used: Excel, VBA

   Responsibility: Was responsible for development of the Reporting Dashboard and generation of a system which will automatically mail daily report to the Sales Managers with different level of visibility, i.e. Center report to center head, Zonal performance report to zonal head, etc.

7. Built a Sentiment Analysis model using Naïve Bayes Classifier to understand the Market sentiment over GST and its compatibility in Tally.

   Overview: After the GST Compatible release of Tally, the Marketing team wanted to understand the running sentiment in the market for the Product and to understand the feature which is most talked about and in a good way or a bad way.

   Technology used: Python, NLTK

   Responsibility: Was responsible for gathering the Twitter Data, Facebook Posts and Comments from Tally help section and putting them in one format and cleaning the Data. Did Several iterations of removing common words, Stemming and Lemmatization of the words were done , bag of words was developed and then a sentiment model was run to rate it into positive or negative sentiment, and word cloud images were developed to understand the most Talked about Features.

8. **Tally Retail:**

   Overview: With Tally Retail, Tally is going to step into the Retail market with its Point of Sale system, CRM, Inventory management, Discount Module, and Retail Analytics, all bundled into one product.

   Technology used: C++, TDL

Responsibility: Analyzing the market requirements, discussing with the stake holders, preparation of high-level designs, coding and debugging. Engaged in rigorous discussion over the Market requirements and High-level designs with the Functional Team. Worked on HLD, LLD and coding of the Product and was closely associated with the Retail Analytics Development through the Design Phase.

9. **Revamping of LIB.BIZ Module:**

    Overview: LIB.BIZ is one of the modules in the Tally.ERP9 Product. This library handles the entire business logic of the product through functions. This library routine supplies the necessary functions to process the data to produce meaningful results.

    Technology used: C++, TDL

    Responsibility: Understanding the changing business requirements to enhance the Language to support the arising needs. Developing high level functional specification documents, designed LLD for the necessary functions and was responsible for development and Debugging, Post release support, Customer issue handling.

10. **Development of E-Banking:**

    Overview: Tally E-Banking is a Module of Tally ERP.9 allowing the customers do his all banking transactions Like Payment, Receipt, Inter-account Transfers through Tally.

    Technology used: C++, TDL

    Responsibility: MY responsibility included Writing Functional Specs and Design Review, Coding and Debugging, Post release support, Customer issue handling. Worked on the Design Analysis and coding of the Various Sub-Features of the module, providing the features of payment instruction generation, upload and Reconciliation of account directly from Tally, in co-ordination with DBS bank.

11. **Development of Service Tax Module:**

    Overview: Tally Statutory Module allows easy creation of Transaction with Tax implications, smooth and quick generation of statutory Reports and facility of Generation of E-returns. It applies the business logic to help the customer find-out if there has been any mistake in the transaction because of which his return filing may go wrong, gives the feature to rectify the mistake on fly and go ahead with the e-return generation.

    Technology used: C++, TDL

    Responsibility: Was deeply engaged in the Functional Flow design to ensure the simplest way for the customer to get compliant with service Tax. My responsibility included understanding the compliance requirements, realizing the pain points of customer, engaging in design and specification document creation and development of the feature in the product.

ACADEMIA

- ≈ B. Tech (EEE), from SRM Univ. Chennai, in 2011   GPA   8.7
- ≈ 12th, from DAV Public School, Ranchi in 2006   Percentage   78.2%
- ≈ 10th, from DAV Public School, Muzaffarpur in 2004   Percentage   86%

COURSES

Coursera:

- ≈ Machine Learning Foundation : A case study Approach
- ≈ Machine Learning: Regression
- ≈ Machine Learning: Classification
- ≈ Machine Learning: Clustering & Retrieval

Personal Details

Name   : Abhinav Kumar
DOB   : 21-Feb-1989
Current City   : Bangalore