**ADITYA GUPTA**

**SUMMARY**

Around 5 years of experience in business solution software and analyzing business operations. Programming experience in Python (such as OpenCV, NumPy, Pandas, Matplotlib). Secured 9 Awards and honors for excellent management of various organizations. Skilled in statistics, analytics, project management and technical and functional reporting, with the strong background in IT, Banking, Telecommunications, Supply Chain and E-commerce retail industries.

# 

**TECHNICAL PROFICIENCIES, SKILLS AND CERTIFICATIONS**

* **Databases:** Oracle, MySQL, SQLite, NoSQL, RDBMS, SQL Server 2018.
* **Project Management:** Waterfall, Agile methodology, Scrum, SDLC
* **Programming Languages:** Python, SQL, C, JAVA, Pyspark
* **DWH BI Tools:** Tableau Desktop, Power BI, MicroStrategy
* **Cloud:** AWS, Azure
* **Statistical Tools:** Regression, Hypothesis Testing, Clustering, VLOOKUP, VBA & Macros, Advanced Excel Functions, Kubernetes
* **Machine Learning:** Supervised and unsupervised learning, Regression, Clustering, SVM, Decision trees, Classification, Recommendation systems,
* **Data Visualization:** Matplotlib, seaborn, ggplot, Tableau
* **Machine Learning Library:** Pandas, NumPy, SciPy, scikit-learn, TensorFlow, spacy

**EDUCATION**

* University of Texas at Dallas, M.S. (I.T. Management) (Outstanding Student Leader Award) August 2016 – May 2018
* Amravati University, B.E. (Electronics and Telecommunications) (Student of the Year Award) August 2009 – May 2013

**PROFESSIONAL EXPERIENCE**

**Technical Analyst/Python Developer,** Advatix, Dallas, Texas, USA July 2018 – November 2018

* **Project:** Business process improvement of a personalized vitamins pack manufacturing company and a baby product manufacturing company.
* **Data Transformation:** Transformed raw data from APIs into MySQL with customized ETL application to prepare unruly data
* **Data Processing and Statistical analysis:** Carried out data processing related to shipping and used statistical techniques such as sampling, estimation, hypothesis testing, time series, correlation and regression analysis using Python as well as Gretl
* **Requirement Gathering and documentation:** Worked with customers, associates and team members to gather functional and non-functional requirements and created SWIs, SOWs and Process Map Visualizations**.**
* **Process Mapping and SOP:** Created various process maps to improve the efficiency of the current operations by analyzing current way of working on various operations. Created SOPs using Dozuki software to document to proper flow in efficient manner.
* **Cloud:** Migrated and implemented multiple applications from on premise to cloud using AWS services.
* **Need Analysis:** Evaluated the technology platforms used by the clients, worked with the clients to understand and document their needs for technology innovation, defined the requirements for new technology platforms, and coordinated with Technology teams in the development and implementation of the technology solutions using machine learning algorithm as classification
* **Report Generation using Python**: Produced detailed analytics and actionable Business Intelligence reports that show key performance indicators, identify areas of improvement and facilitate informed decision-making using python libraries as Seaborn and matplotlib
* **Machine learning algorithms used**: Classification, KNN, Linear regression, Logistic regression, Decision tree
* **Technologies**: Python, Tableau, MySQL, Debugging, ETL, GRETL, NLTK, Pyspark.
* Excellent knowledge of **Machine Learning**, Mathematical Modeling and Operations Research along with R, Python, MATLAB, and Relational databases. Deep understanding & exposure of Big Data Eco-system.
* A highly immersive Data science program involving data Manipulation & Visualization, Web Scraping, **Machine Learning**, Python programming, SQL, GIT, MongoDB

**Data Analyst -** Intelesystems, Dallas, Texas, USA June 2017 – August 2017

* **Project:** Designed an exponential smoothing model for AT&T to iteratively forecast future values of a regular time series of values from weighted averages of past values of the series using Tableau
* **Data Visualization and Business Prediction:** Visualized and created a unique model of various revenue forecasting reports by using Tableau for international billing information of AT&T for customers in roaming zones
* **Clustering Analysis:** Conducted Clustering analysis to generate segmented profiles of customers to analyze targeted areas, carriers and customer expense trends using machine learning algorithms
* **Technologies:** Excel, Tableau, Data Analysis Plus, Python, PyCharm (IDE)

**ACADEMIC AND RESEARCH PROJECTS** August 2016 – June 2017

* **Webcam Motion Detector:** This is the program that detects moving objects in front of a computer webcam and records the time that the object enters the webcam, the video frame and the time when the object exited the video frame. It is build using Python, image processing and video processing
* **Scraping of Real Estate Property Data from the Web:** This project scrapes property data from a real estate websites, yellow pages or information from Wikipedia using web scrapping using Python. It gets the data and store them in a table i.e. Pandas data frame. To do web scrubbing on this real estate website with Python, it performs a search query on a certain location so Python will search for data and then it gets that data, and stores it in the table, and then finally in an output CSV file, so everything is automatic
* **Desktop Application to store information of Books:** Developed a desktop graphical program with a database backend program that allows the user to store information about books. Program is an executable file that run on windows and dot app files that run on Mac. It is built mainly with Tkinter which a graphical user interface library and the SQLite3 library which is a library to interact with SQLite database
* **Interactive Dictionary using Python:** Created an interactive dictionary which takes Json data file and extracts definition of word along with identification of case sensitivity, suggesting similarity of two words if there is spelling error, recommending best match and user confirmation

**Sr. System Analyst,** Techrel Technologies, Pune, Maharashtra, India February 2013 – June 2016

* **Project:** To identify the language of the file it is written in when file extensions are not written properly, or file extensions are missing for a client moving existing data to Aws cloud to be stored in formatted manner to gather static metrics for Java, JavaScript, & config files
* **Challenge:** To identify language of the file when developer uses a non-standard extension such as name of component it relates to or doesn’t add file extension at all to know which parser should be used to gather the static metrics.
* **Data Scrapping:** Created Pandas dataframes by reading file lists into data frames then sampled them using reproducible seed by using library as random, os for os\_walk, pattern matching using fnmatch
* **Predictive model using Machine learning algorithm:** Created a prediction target and explored to target data using sklearn. Used number tokenization method to isolate numbers from words and symbols using nltk. Identified frequency of most common words in file using frequency distribution function in nltk. Visualized the top 30 characters appeared in tokenized frequency using Seaborn library of python. The tokenizing of texts and building a bi-directional LSTM (Deep learning model), with single- class prediction helped in predicting probability if the code is of particular language or not.
* **Training the model:** Initialized the vectorizer and created sparse feature matrix to transform words into numerical tokens to be used for modelling using Tfidfvectorizer. Trained Naïve Bayes model on training data where input to the model was vectorized data and output was probability if the file was of particular language or not.
* **Visualizing the model performance:** Created a confusion matrix to identify test predictions to check the accuracy by finding true positives and false negatives. Visualized final out put using histogram to check whether the prediction is correct or not. Confusion matrix visualization helped in demonstrating how well the model performed
* **Libraries used:** OS, Fnmatch, random, pandas, matplotlib, sklearn, numpy, nltk, seaborn, urllib
* **Languages extracted:** Java, C, C#, Python
* **IDEs used:** Pycharm, Jupyter, Anaconda, Atom
* **Project 2:** Fraud detection of Credit transactions and identification of customer satisfaction
* **Prediction Model:** Designed and developed a prediction models for new areas to be explored for new customer segment and optimizing market behaviors while introducing new banking products in the target market using linear regression algorithm for behavior analysis and logistic regression for new product introduction to check satisfaction reports and check fraudulent transactions using classification algorithm of machine learning
* **Data structure analysis:** Analyzed the existing data, created new data structures to identify transactional records to examine customer trends which further helped in predicting customer satisfaction rate using Tensorlflow
* **Sales and Marketing Analysis using Python:** Worked with sales and Marketing team for collaborating with cross functional team to frame and answer important questions using data analysis with the help of Matplotlib and Seaborn libraries in Python
* **Data Model Design and Integration:** Designed a data model for a multi-vendor product environment of ‘Gramin’/ Rural Banking sector, being used for systems integration, operational data storage, reports, and data migration
* **Business Analysis:** Ensured 100% alignment of deliverables with business requirements of ‘ATM Switching’ software, solution design and integration design by collaborating with vendors and technical delivery team
* **Technologies:** Java, Oracle SQL, DB2, Python, Excel, NLP, AWS, Powershell
* **Python and machine learning libraries:** OS, PySpark, Tensorflow,

**VOLUNTEER EXPERIENCE**

**PRESIDENT,** ENACTUS, The University of Texas at Dallas, January 2017- May 2018

* **Award Winning Leadership:** Secured 5 awards from Enactus and the University including fellowship from Bank of New York, Student Leader of the Year award and Diamond Service leadership for serving more than 1500 hours
* **Establishment and Reformation of Organization:** Developed a new organizational structure from zero to 100 members for UT Dallas and helped in increasing team productivity by 12 times by starting 5 CSR projects in one year.

**ACHIEVEMENTS**

* Outstanding Student leader of the year (2018), BNY-ENACTUS student fellow (2018), Enactus Diamond service leadership (2018)
* Enactus Platinum service leadership award (2017), Enactus Service Leadership Award (2017)
* Student of the Year (2013), Influential Young Leader award (2013), Young change Maker (2011)