

DIMPLE BAPNA

bapna.d@northeastern.edu | [Git Hub](#) | [Linked In](#) | [Portfolio](#) | Seattle, WA | (425) 233-7801

SUMMARY

Data Scientist with a growth mindset and a passion to dive deep into data and derive strategic insights that impact business and customers. Strong communication skills with hands-on practice of problem-solving to deliver business value.

EDUCATION

Northeastern University, Master of Science, Data Analytics Engineering **GPA: 3.84/4** Aug 2019 - Dec 2020

Courses: Data Mining, Neural Nets and Deep Learning, Data Warehousing and Business Intelligence, Database Management & Design, Data Visualization, Probability & Statistics, Deterministic Operations Research, Statistical Methods in Engineering

University of Mumbai, Bachelor of Engineering, Computer Engineering Aug 2008 - Jun 2012

Relevant courses: Data Structures & Algorithms, Object Oriented Programming (Java), Mathematics, Artificial Intelligence

EXPERIENCE

Data Scientist - Consultant **Costco Wholesale, Seattle** Jun 2020 - Aug 2020

Tech stack: Python, pandas, scikit-learn, matplotlib, seaborn, neural network (LSTM), Tableau

- Implemented LSTM model which forecasts demand of goods sold by Costco.com, reducing revenue loss by 20%
- Developed comprehensive data presentation reporting forecasted demand at depot-level using Tableau dashboard
- Leveraged data exploration on large data sets to identify limitations and recommended improvements in data quality

Data Scientist - Consultant **Tignis, Seattle** Feb 2020 - Jun 2020

Tech stack: Python, pandas, scikit-learn, matplotlib, neural network

- Implemented anomaly detection for real-time sensor data with neural networks (Feed forward) by analyzing data
- Performed statistical modeling on sensor data for data manipulation and missing value treatment
- Reduced the overall time taken for fault investigation by 60%, resulting in direct cost savings of more than \$300k

Senior Software Developer **Duck Creek technologies, India** Aug 2016 - Jun 2018

Tech stack: C#, SQL, Duckcreek tools

- Led a team to create and manage an end-to-end billing system resulting in a 35% faster billing process
- Collaborated with cross-functional teams and business stakeholders to articulate a data-driven strategy to improve billing platform; reducing customer attrition rate by 20% and enhancing customer experience
- Developed an automated payment records system compatible with 50+ banks, saving 160 man-hours per month

Software Developer **Accenture Services Pvt Ltd, India** Feb 2013 - Jul 2016

Tech stack: C#, SQL

- Automated medical finance application for incident analysis, increasing productivity of the healthcare system by 70%
- Developed a single input application which eliminated manual interventions reducing engineering costs by over 60%
- Delivered a tool to eliminate duplicates and highlight pending invoices from financial data saving \$2M+ in interest penalties for the client

DATA SCIENCE PROJECTS

Case Study: Customer acquisition for Nike using Customer Analytics <https://tinyurl.com/NikeByYou> Dec 2020

- Performed customer segmentation on existing Nike customers using marketing analytics and ML model on customer data, sales history, etc.
- Targeted customers for personalized recommendations using Market Basket analysis and sales history
- Evaluated the monetary value of the recommended strategy using Customer Lifetime Value Financial analysis

Stock Exchange Data Warehousing and Business Intelligence Dec 2020

Tech stack: AWS lambda, S3, Python, SSIS, SQL Server (Relational Database), SSMS

- Designed ETL pipeline and performed data modeling using SSIS to integrate data into facts and dimensions
- Automated extraction of data from multiple data sources using APIs, AWS lambda and exported it to S3 bucket
- Accomplished data collection, profiling, cleansing, integration and price history tracking using SSIS

Seoul Bike Sharing Demand using 2 factor experimentation Dec 2020

Tech stack: SAS, Python

- Performed statistical analysis of variance (ANOVA) of a two-factor experiment to study the effects of Seasons and Holiday/ No-Holiday day on the number of bikes rented, and if there is a significant interaction between the two factors
- Conducted Tukey test to further group the factors and find the groups that differ significantly from each other

Research Thesis: Image classification using Convolutional Neural Networks Oct 2020

Tech stack: CNN, TensorFlow, Python

- Designed a Convolutional Neural Network for image classification using a large data set CIFAR-100
- Improved the accuracy of the model using various optimization techniques like regularization, image augmentation, etc.

Machine learning based SMS message filter Dec 2019

Tech stack: Python, Sklearn, NLP, Multinomial Naive Bayes algorithm

- Processed dataset of raw SMS messages using data analysis and data wrangling and performing feature selection
- Applied Natural Language Processing and Multinomial Naive Bayes machine learning algorithm to create a spam detection filter to screen out unwanted and abusive messages for a user

TECHNICAL SKILLS

Programming Skills: Python (Pandas, NumPy, SciKit-learn, matplotlib, TensorFlow, Keras, PyTorch, Postgres)

Machine Learning Algorithms : Linear Regression, Logistic Regression, Decision Trees and Random Forests, KNN classifier, K-Mean Clustering, SVM, Apriori, Neural Nets and Deep Learning

Data Science / Analytics: R, SAS, SQL, Tableau, SSIS, SSAS, Excel

Other Applications: Version Control Systems (Git, Bitbucket), GitHub, Jupyter Notebook, IntelliJ, JIRA, Confluence, Lingo