ASHIR MEHTA

Boston, MA | +1(857) 334-7053 | Linkedin | mehta.ash@northeastern.edu

EDUCATION

Northeastern University, Khoury College of Computer Sciences — Boston, MA
Master of Science in Data Science**,** GPA: 4.0 / 4.0 — Expected May 2023
Coursework: Introduction to Data Management & Processing, Intro to Programming for Data Science

Indus University — Ahmedabad, India
Bachelor of Technology in Computer Engineering, GPA: 9.14/10.00 — Aug 2017 - May 2021
Coursework: Big Data Analytics, IOT, Python, Data Warehouse & Mining, Design and Analysis of Algorithms, Web Technologies, Software Engineering, Data Structure, DBMS, C++, C.

SKILLS & TECHNOLOGIES

Python, Numpy, Pandas, Scikit-Learn, Tensorflow, Matplotlib, Seaborn, NLTK, R, SQL, Tableau, MS Excel, Hadoop, Selenium, Web Scraping, Javascript, Probability, Statistical Data Analysis, Data Visualization, Data Manipulation, Predictive Modelling, Time Series Analysis, Deep Learning, Machine Learning, MS-PowerPoint, C++, C, Git, HTML, CSS, MATLAB, Slack, Jira, LaTeX, Linux

WORK EXPERIENCES

Sidhhi.ai — Bengaluru, India
Data Extraction Engineering Intern — Jan 2021 - May 2021
- Built 50 scalable web crawlers in selenium and python to extract 2.5 million customer reviews from 26 fashion ecommerce websites including Sephora, Nordstrom, Ulta, Amazon, Walmart and social media platforms such as Youtube.
- Established data cleaning, data pre-processing, and data manipulation on ~3 GB of CSV data files to be utilized to train deep learning models for performing sentiment analysis and people analytics.
- Implemented statistical data analysis and visualization to derive practical insights on customer behaviour trends present in dataset to grow client's revenue and popularity amongst consumers.

Exposys Data Labs — Bengaluru, India
Data Science Intern — Nov 2020 - Jan 2021
- Analyzed gathered customer data by administering an unsupervised machine learning algorithm of k-means clustering and statistical techniques to investigate shopping habits of customers.
- Performed user segmentation based on annual income, spending score, and age to identify and target potential customers to increase sales of shopping malls by 20%.
- Coached 10 junior engineers in Python, R, SQL and Tableau concepts to develop data visualization and problem-solving abilities.

Cppsecrets Technologies — Noida, India
Software Developer Intern — Sep 2020 - Nov 2020
- Collaborated with 5 developers to accomplish 2 organizational projects on chatbot development and time series analysis to reduce overall development time by 2 hours.
- Organized, presented, communicated, and reported results to project heads through PowerPoint presentations, Slack, and Jira.
- Cataloged and published 26 technical articles on machine learning and python programming theories for company's website.

PROJECTS

Northeastern University, Wikipedia Web Traffic Analysis & Time-Series-Forecasting — Dec 2021
- Evaluated ML models of XGBoost, ARIMA, LSTM, and Prophet to carry out web traffic time series forecasting of wikipedia pages.
- Executed exploratory data analysis and visualization with 5 team members in R programming to find trends and outliers in data.
- Concluded that LSTM and Prophet models were most appropriate models for web traffic analysis based on mean absolute error values.

Boston University, WellBud — Nov 2021
- Designed a system to address mental health issues in a 27 hours long hackathon by utilizing twitter's english messages dataset.
- Conducted natural language processing techniques to format ~20,000 row entries of data to construct deep learning model for classifying user's sentiment as either of joy, sadness, fear, anger, surprise, or love with 93.05% accuracy.
- Improvised system by creating basic chatbot to calm users whose sentiment is predicted to be either of sadness, fear, or anger.

Cppsecrets Technologies, Instagram Engagement Analysis — Nov 2020
- Developed an instagram bot utilizing python, page object pattern, and selenium to automate account login process, like, comment and send follow requests to 100 enlisted instagram profiles per day.
- Integrated time series forecasting with ARIMA and SARIMA to compute proportionality of people following back per day and monitor marketing strategy success rate.

Indus University, Predicting CO2 Emission using Multivariate Linear Regression — Sep 2020
- Programmed multiple linear regression model to estimate CO2 emission from vehicles based on 6 independent features.
- Achieved 0.8481 r2 score, 0.90 variance score, and 435.56 Residual sum of squares for model using Scikit-Learn.

PUBLICATIONS

- **Denouements of machine learning and multimodal diagnostic classification of Alzheimer's disease**
  Review Paper Published in Visual Computing for Industry, Biomedicine, and Art Journal of Springer — Nov 2019 - Nov 2020

- **The impacts of artificial intelligence techniques in augmentation of cybersecurity: a comprehensive review**
  Review Paper Published in Complex & Intelligent Systems Journal of Springer — Sep 2019 - Mar 2020