Aditya Sanjay Mhaske

admhaske@iu.edu | LinkedIn/adityamhaske | Github | adityamhaske.com

EDUCATION

Indiana University Bloomington — Bloomington, IN, USA
Master of Science - **Data Science** — May 2024
Coursework: Machine Learning, Statistics, Probability, AI, Big Data, Data Visualization, and Economics II

MIT World Peace University — Pune, India
Bachelor of Technology - **Computer Science** — Jun 2022
Coursework: Data Warehousing, Database Management, Data Mining, Project Management, and Software Design

SKILLS

Languages: Python (PySpark, Pandas, NumPy), R, SQL (MySQL, PostgreSQL), Matlab, HTML, CSS, MongoDB
Big Data: PowerBI, Tableau, Excel, Azure, Airflow, AWS, Data Bricks, Git, Google Analytics, Cuda, Neo4J, SAS
MLOps: Docker, FastAPI, Flask, Transformers, ETL, Snowflake, REST, Hugging Face, LLM

WORK EXPERIENCE

Data Scientist | Kelley School of Business — Dec 2022 - May 2024
- Implemented **Llama 2** with meta 7B and 13B in an **LLM Model** to analyze Customer Satisfaction and Brand Equity for large-size US companies, enabling the prediction of binary classification percentages for two categories.
- Leveraged BERT **NLP** algorithm to analyze a large-scale dataset of **100+ million rows and 40 columns** and attained testing accuracy of 91% in classifying topics and sentiments of political campaigning data.
- Created a predictive Cross-Classified Multilevel model, boosting decision-making and performance by 35% through data analysis and pattern identification.

Data Scientist Intern | Twin Cities Innovation Alliance — Sep 2023 - Dec 2023
- Designed **A/B tests** and **Hypothesis tests**, employing SQL for data preparation and **Statistical Modeling** like Segmentation and Regression analysis to assess trends and evaluate the effects of changes on website optimization.
- Improved client's recommendation system by blending collaborative and content-based filtering techniques, utilizing user preference and behavior data, leading to a notable **20% surge** in user engagement acquisition.

Data Engineer | Moonplexus Private Limited — May 2021 - Aug 2022
- Led the AWS team to implement **AWS EC2** and **Amazon RDS**, for database management. Established an ETL pipeline to gather images, ensuring data integrity and optimizing data storage for efficient retrieval and Analysis.
- Developed machine learning models for skin lesion classification using **TensorFlow** and **CNN,** achieving **89%** accuracy.
- Implemented production-ready machine learning model with **FastAPI, Docker,** and **Flask**, resulting in a streamlined deployment process, increased scalability, and enhanced user accessibility.
- Prioritized SQL integration to automate tasks, and optimize data retrieval, achieving a 40% reduction in execution time.

Machine Learning Intern | Indian Institute of Technology — Dec 2022 - Aug 2023
- Employed Python and C++ to analyze DBSCAN and MBSCAN algorithms to analyze distance-based outlier mining. Improved computational **speed by 30%** and **precision by 20%** in clustering algorithms.
- Incorporated Data Modeling for feature engineering, MS Excel for data preprocessing, and Tableau for visualization, enhancing the understanding and presentation of insights derived from the analysis of results.

ACADEMIC PROJECTS

Stock Prediction on Deutsche Borse using AWS (Link) | Python, AWS, Time Series — Jan 2023 - Mar 2023
- Forecasted EUR currency start prices using Python to construct and deploy ARCH and GARCH models.
- Employed AWS EMR and Sagemaker to seamlessly consolidate and process a vast dataset of over 1000 CSV files, totaling 45.5 million data points stored in an S3 bucket.

Multi-modal for Depression Analysis (Link) | Python, NLP, Feature Engineering — Jun 2022 - Dec 2022
- Developed a fusion-based multimodal approach incorporating EEG signals, speech data, and facial features to develop an accurate (90%), and efficient system for early-stage clinical depression detection. (IEEE Publication)
- Leveraged Python for feature extraction, and data augmentation with Machine Learning and Deep Learning algorithms.

ACHIEVEMENTS

- Secured the 6th position and led the team in the AGBI Health-Tech Challenge, among 4,000 participants.
- IEEE Publications and Patents: Object Detection and Localization (Computer Vision), and Cardiovascular Disease Predictions.