PRANAY BHAKTHULA
Washington, DC 20037 | pranaybhakthula@gmail.com | (202) 469-2775 | Linkedin | Github | Portfolio

EDUCATION

The George Washington University (GWU), Columbian College of Arts and Science Washington DC, United States
Data Science Masters Candidate (CGPA: 3.95*)* December 2022
- Relevant Coursework: Deep Learning, NLP, Time Series Analysis, Data Mining, Data Warehousing, Cloud Computing

Sathyabama Institute of Science and Technology Chennai, India
Bachelor of Engineering, Electronics and Communication (CGPA: 8.66/10) May 2019

TECHNICAL SKILLS

- **Programming languages**: Python, SQL, R, C/C++, Scala, JavaScript, HTML, CSS
- **Database**: MySQL, Microsoft SQL Server, PostgreSQL, NoSQL- MongoDB, DynamoDB
- **Python Packages**: Numpy, Pandas, Scikit-Learn, Scipy, Keras, Tensorflow, PyTorch, Matplotlib,
- **Machine Learning**: Clustering, Regression, Classifiers, Natural Language Processing (NLP), Image Classification
- **Tools**: Excel, Tableau, Microsoft Power BI, SPSS, SAS, SSMS, SSIS
- **Cloud technologies**: Amazon Web Services (AWS), Google Cloud Platform (GCP), Microsoft Azure, Snowflake, Alteryx

RELEVANT WORK EXPERIENCE

Solutions Architect Intern Seattle, United States
Amazon Web Services (AWS) May 2022 – August 2022
- Implemented real time video analysis of basketball shot to provide shot and posture analysis of the shooter with accuracy of 76% by object detection and image classification using Tensorflow, OpenCV, CNN, Yolov7 models
- Designed and tested 8 different algorithms using AWS S3, AWS EC2 to find the most optimized model, resulted in increasing the accuracy by 15% and decreasing the runtime by 25%
- Learned AWS sales procedures/methods and tools available to assist customers by shadowing meetings between Solutions Architects and 4 different AWS customers

Student Administrative Assistant Washington DC, United States
George Washington School of Public Health (GWSPH) November 2021 – December 2022
- Collaborated with IT team to implement new data management and reporting systems using SQL Server and Tableau, resulting in a 70% reduction in report generation time
- Developed and maintained SQL queries to extract and update research expenditures of departments, exported the results as Excel sheets from SSMS, which reduced analyzing time by 40%
- Collaborated in creating monthly reports of expenditures by developing pivot tables, graphs and dashboards using Tableau and Excel, resulted in 30% increase in having productive meetings
- Automated data entry by extracting information from PDF invoices using Python, resulting in an 80% reduction in data entry time and exporting relevant data to Excel sheets

Data Analyst Andhra Pradesh, India
Centre for Rural Studies and Development (CRSD) June 2019 - July 2021
- Filtered Federal and State budgets data in departments such as Agriculture, Education, Health using SQL queries in SSMS and exported required data in Excel sheets, which led to 50% reduction in analyzing time
- Automated ETL (Extract, Transform, Load) process using SSIS from heterogenous data sources which led to 60% reduction in reporting time
- Created monthly reports by using Pivot tables, filters, functions in Excel and visualization tables, graphs and dashboards using Tableau, which increased campaigning by 150% for state level policy change
- Automated web scraping using Python to collect data for monthly and annual reports, resulting in 100 hours reduction in manual data collection time
- Maintained data entry of field reports and audited financial expenses of the organization using Filters, Pivot tables in Excel to find discrepancies and minimized the expense by 8%

TECHNICAL PROJECTS

Anomaly Detection in Wood fossil | Python November 2022
- Implemented a method to detect the different patterns in wood fossil images to classify damaged and undamaged part of the wood with an accuracy of 86% by applying Resnet50, VGG19 and PCA models

Fake or Real News Classification | Python April 2022
- Classified news as real or fake with an f1-score of 0.9916 using Natural Language Processing (NLP) algorithms such as DeBERTa, RoBERTa, DistilBERT, with RoBERTa having highest f1-score
- Trained the models using AWS EC2 and GCP compute engines which reduced the runtime by 80%

Cotton Plant Disease Prediction | Python December 2021
- Predicted diseased plants with 96.8% accuracy by training 12K images with CNN and pre-trained models (Resnet50, VGG16, Densenet121) on AWS EC2 and GCP compute engines

Kobe Bryant Shot Selection | Python June 2021
- Predicted shots on Bryant's career with 67% accuracy by building MLP classifier and Random Forest classifier
- Determined strengths and weakness from opponent's perspective in challenging Bryant in the 2 out of 5 finals by conducting EDA analysis